

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 624 Seconds
(without alignments)
8033.424 Million cell updates/sec

Title: US-10-000-151b-1

Perfect score: 1857
Sequence: 1 atggcgcgcgtatagaggtta.....aagtaatatgncncaaa 1857

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_19Jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1810.6	97.5	1857	24	ABK86572
2	1645.4	88.6	2280	24	AAAD44094
3	1359.8	73.2	1461	24	ABT05463
4	1311.8	70.6	1489	24	ABT05462
5	1217.2	65.5	2362	22	AAH18539
6	899.6	48.4	1165	21	AAA87668
7	894.6	48.2	1160	21	AAA87699
8	725.2	39.1	807	22	AAH06644

9	509.4	27.4	843	24	ABZ11754
10	502	27.0	539	21	AAZ58638
11	350.4	18.9	387	22	AAK89423
12	343.2	18.5	469	22	AAFP93488
13	336.4	18.1	442	22	AAFP93569
14	334.2	18.0	348	22	AAH98871
15	258.4	13.9	283	20	AAV87898
16	174.6	9.4	856	22	AAH06371
17	174.6	9.4	2245	22	AAH16194
18	151.8	8.2	546	22	AAH11428
19	86.8	4.7	10233	23	ABL13164
20	85.2	4.6	3672	23	ABL13165
21	72	3.9	6644	20	AAK33181
22	72	3.9	7372	20	AAK33182
23	72	3.9	7797	20	AAK33180
24	72	3.9	7996	20	AAK33184
25	71.2	3.8	4545	23	ABL10718
26	71	3.8	6668	24	ABU3697
27	68	3.7	14006	24	ABL33958
28	67.2	3.6	6292	22	AAK46735
29	65.8	3.5	16033	24	ABL34404
30	65.2	3.5	8079	24	ABL92313
31	65	3.5	3683	25	ABZ10199
32	65	3.5	7442	22	AAK46686
33	64.8	3.5	9155	24	ABL32462
34	64.2	3.5	6898	24	ABN80222
35	64.2	3.5	9539	22	AAK45347
36	64.2	3.5	9539	24	ABK28180
37	63.6	3.4	6767	22	AAK46608
38	63.4	3.4	34769	22	AAK46774
39	63	3.4	875	22	AAK46781
40	62.6	3.4	4654	22	AAK46781
41	62.6	3.4	4654	24	ABL34223
42	62.2	3.3	7571	24	ABL32527
43	62	3.3	17934	24	ABL33719
44	61.6	3.3	819	22	AAK46774
45	61.6	3.3	863	22	AAK46774

ALIGNMENTS

RESULT 1	ABK86572	standard; CDNA; 1857 bp.
ID	ABK86572	
XX	ABK86572;	
XX		
DT	24-SEP-2002 (first entry)	
XX		
DE	CDNA encoding human ether-a-go-go related interacting protein KCRIb.	
XX		
KW	Human; ss; gene; human ether-a-go-go related gene; HERG; KCR1; SNP;	
KW	Long QT syndrome; LQT; single nucleotide polymorphism; cardiac arrhythmia;	
KW	potassium channel.	
XX		
OS	Homo sapiens.	
XX		
PH	key	Location/Qualifiers
FT	CDS	1..1422
FT		/tag a
FT		/product- "KCRIb"
FT	variation	replace (1339,A)
FT		/tag b
FT		/standard_name- "Single nucleotide polymorphism"
XX		
PN	W0200242735-A2.	
XX		
PD	30-MAY-2002.	
XX		
XX	30-OCT-2001; 2001WO-US45644.	
XX		
PR	30-OCT-2000; 2000US-244340P.	

XX (UYVA-) UNIV VANDERBILT.
XX
XX
XX Balser JR, George AL, Roden DM;
XX
XX MPI: 2002-527650/56.
XX P-PSDB; AA099166.
XX
XX Identifying a potassium channel activity modulator for drug design,
XX comprises contacting a compound with a potassium channel and rat
XX cerebellar cDNA library (KCR1) polypeptide, and determining activity
XX
XX
XX Claim 11: Page 154-156; 164pp; English.
XX
XX The invention relates to identifying (M1) a compound that modulates
XX biological activity of a potassium channel (PC), by contacting a
XX compound with a structure comprising a PC polypeptide and a polypeptide
XX cloned from a rat cerebellar cDNA library (KCR1), and determining the
XX activity of the PC polypeptide in the presence and absence of the
XX compound, where a difference in the activities indicates modulation of
XX biological activity of PC. Also included are identifying (M2) a candidate
XX compound that modulates the biological activity of a complex comprising a
XX human ether-a-go-go-related gene (HERG) channel polypeptide and a KCR1
XX polypeptide, identifying (M3) a candidate compound as a modulator of KCR1
XX expression, modulating (M4) PC function in a subject, comprising
XX administering to the subject a substance that provides expression of a
XX KCR1-encoding nucleic acid molecule in a cell or tissue, where modulated
XX PC function is desired, screening (M5) for susceptibility to a drug-
XX induced cardiac arrhythmia in a subject, comprising obtaining a
XX biological sample from the subject and detecting a polymorphism of a KCR1
XX gene in the biological sample from the subject, where the presence of the
XX polymorphism indicates the susceptibility of the subject to a
XX drug-induced cardiac arrhythmia, an oligonucleotide pair, where a first
XX oligonucleotide of the pair hybridizes to a first portion of a KCR1 gene
XX which includes a polymorphism of the KCR1 gene, and the second
XX oligonucleotide of the pair hybridizes to a second portion of the KCR1
XX gene that is adjacent to the first portion and a set of antisense
XX oligonucleotide primers, suitable for amplifying a portion of a KCR1 gene
XX which includes a polymorphism of the KCR1 gene. (M1) is useful for
XX identifying a compound that modulates biological activity of PC,
XX especially HERG, for modulating PC function (i.e. modulating HERG
XX activity) in a mammal, by preparing a composition comprising the
XX compound and administering the composition. The compound is useful for
XX treating or preventing long QT syndrome (LQT) and is useful in drug
XX designing. The present sequence encodes the human KCR1b allele.
XX
XX
XX Sequence 1857 BP; 475 A; 343 C; 334 G; 652 T; 53 other;
XX
XX
XX Query Match 97.5%; Score 1810.6; DB 24; Length 1857;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1857; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 1 ATGGCGCAGTAGAGGTTACTGTTTCGCGCGCTGAGCTGACCTTTTATAGTCC 60
XX 1 ATGGCGCAGTAGAGGTTACTGTTTCGCGCGCTGAGCTGACCTTTTATAGTCC 60
XX
XX 61 TGCCTCTCTCTCTCGCGCTTACGCCGGGCGTGGGAGAGCCCTACATGAGAGATCTTC 120
XX 61 TGCCTCTCTCTCTCGCGCTTACGCCGGGCGTGGGAGAGCCCTACATGAGAGATCTTC 120
XX
XX 121 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCGCATTTCTCCCTTCCCATGGGATCC 180
XX 121 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCGCATTTCTCCCTTCCCATGGGATCC 180
XX
XX 121 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCGCATTTCTCCCTTCCCATGGGATCC 180
XX 121 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCGCATTTCTCCCTTCCCATGGGATCC 180
XX
XX 181 ATGATTAATACATTACCTGCTTACCTGCTGAGTGGAGTGGTCAACCTGCCATT 240
XX 181 ATGATTAATACATTACCTGCTTACCTGCTGAGTGGAGTGGTCAACCTGCCATT 240
XX
XX 241 TGGATCTTTGGATGCTGAGACATGTTGCTGCTCATTTGGAGTGCATGATTTGTTAAT 300
XX 241 TGGATCTTTGGATGCTGAGACATGTTGCTGCTCATTTGGAGTGCATGATTTGTTAAT 300
XX
XX 301 CTTCCTTTCAGGTGGCAACTTCTATTACTATTGTTTGTTCACAGAGTACAAACC 360
XX 301 CTTCCTTTCAGGTGGCAACTTCTATTACTATTGTTTGTTCACAGAGTACAAACC 360

DB 301 CTTCCTTTCAGGTGGCAACTTCTATTACTATTGTTTGTTCACAGAGTACAAACC 360
DB 361 AGAACAAGGCTGCGCTCAAGTATCAGAGAGTCTTGTCAATTAACCTAGCAGTATTT 420
DB 361 AGAACAAGGCTGCGCTCAAGTATCAGAGAGTCTTGTCAATTAACCTAGCAGTATTT 420
DB 421 CCAACACTTTATTTTAACTCCCTTATTAACAGAGAGATCTATGTTTAACT 480
DB 421 CCAACACTTTATTTTAACTCCCTTATTAACAGAGAGATCTATGTTTAACT 480
DB 421 CCAACACTTTATTTTAACTCCCTTATTAACAGAGAGATCTATGTTTAACT 480
DB 421 CCAACACTTTATTTTAACTCCCTTATTAACAGAGAGATCTATGTTTAACT 480
DB 481 CTTTGTGATATTTGATGCTCTTTATGGAATATATAAATCTAGCTTCTTGATTT 540
DB 481 CTTTGTGATATTTGATGCTCTTTATGGAATATATAAATCTAGCTTCTTGATTT 540
DB 541 TGTGGCTTCATGTTTGGCAACAATATCATCTGGGCTGCTCTGCGAGGAAATGTC 600
DB 541 TGTGGCTTCATGTTTGGCAACAATATCATCTGGGCTGCTCTGCGAGGAAATGTC 600
DB 541 TGTGGCTTCATGTTTGGCAACAATATCATCTGGGCTGCTCTGCGAGGAAATGTC 600
DB 601 ATTCGACAAAAGTTAACTGAGGCTTGGAAAATGAGCTACAAAAGAGAGACAGACTT 660
DB 601 ATTCGACAAAAGTTAACTGAGGCTTGGAAAATGAGCTACAAAAGAGAGACAGACTT 660
DB 661 CCACCTATTAAGAGACATTTGGCAAAATTCGAAAATCTTCAGTCTTTGGCTTAT 720
DB 661 CCACCTATTAAGAGACATTTGGCAAAATTCGAAAATCTTCAGTCTTTGGCTTAT 720
DB 721 TCCATATGCTTTAAACCTTGAATGATGCTTTCTGTTGACTTTGGCCCTACATCTCTG 780
DB 721 TCCATATGCTTTAAACCTTGAATGATGCTTTCTGTTGACTTTGGCCCTACATCTCTG 780
DB 781 GGAATTTGTTTGTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GGAATTTGTTTGTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 840
DB 841 AGTATGAGAGGCTGCTTCAATTTCTCAATATTTCTTCTTCTTCTTCTTCTTCTTCT 900
DB 841 AGTATGAGAGGCTGCTTCAATTTCTCAATATTTCTTCTTCTTCTTCTTCTTCTTCT 900
DB 901 TTTTCTTTCTCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
DB 901 TTTTCTTTCTCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
DB 961 AAACATGGAATTCGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020
DB 961 AAACATGGAATTCGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020
DB 1021 ACTTATGCTCAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
DB 1021 ACTTATGCTCAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
DB 1081 AGAGTTTTCGAAATATGCAATTCGTAATTAATTTAGTTCCAGCCTATATATTTGCT 1140
DB 1081 AGAGTTTTCGAAATATGCAATTCGTAATTAATTTAGTTCCAGCCTATATATTTGCT 1140
DB 1141 GGTGAGATATAGCTGACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
DB 1141 GGTGAGATATAGCTGACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
DB 1201 ATATGCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
DB 1201 ATATGCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
DB 1261 CTTATGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1320
DB 1261 CTTATGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1320
DB 1321 CTGAGTGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
DB 1321 CTGAGTGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
DB 1381 CAGTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1440
DB 1381 CAGTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1440

QY 601 ATTGCACAAAAGTTAACTGAGCTTGGAAAACTGACTACAAAGAGAGACAGACTT 660
 Db 694 ATTGCACAAAAGTTAAAGGAGAGCTTGGAAAACTGAGCTACAAAGAGAGACAGACTT 753
 QY 661 CCACCTATTAAAGAGACCTTTGCGAATTCAGAAAAATTCCTCAGTTCTTTGGCTTAT 720
 Db 754 CCACCTATTAAAGAGACCTTTGCGAATTCAGAAAAATTCCTCAGTTCTTTGGCTTAT 813
 QY 721 TCCATGTCCTTTAAAACTGAGTATGCTTTCTGTTGACTTGGCCCTACCTCTCTG 780
 Db 814 TCCATGTCCTTTAAAACTGAGTATGCTTTCTGTTGACTTGGCCCTACCTCTCTG 873
 QY 781 GGATTTCTGTTTGGCTTTGTTGTTAGTAAATGGTGAATTTGTTTGGCATCGAGT 840
 Db 874 GGATTTCTGTTTGGCTTTGTTGTTAGTAAATGGTGAATTTGTTTGGCATCGAGT 933
 QY 841 AGTCATAGAGCCCTGCTTCACTTTTCCCACTATTCACCTTTTTCATTTACCTCTCTT 900
 Db 934 AGTCATAGAGCCCTGCTTCACTTTTCCCACTATTCACCTTTTTCATTTACCTCTCTT 993
 QY 901 TTTTCTTCTCCTACCTCCTGCTCCTAGCAAAATTAAGACTTTTCTTCTTCTTGGT 960
 Db 994 TTTTCTTCTCCTACCTCCTGCTCCTAGCAAAATTAAGACTTTTCTTCTTCTTGGT 1053
 QY 961 AAACATGGAATTCCTGTTTGGTGGTTACCTTACTCTGTTTGGTTTGAATTC 1020
 Db 1054 AAACATGGAATTCCTGTTTGGTGGTTACCTTACTCTGTTTGGTTTGAATTC 1113
 QY 1021 ACTTATGCTCAATTAATCTTGTCTAGCAGACAAATAGACATTAATCTTCTATGTGTGAAA 1080
 Db 1114 ACTTATGCTCAATTAATCTTGTCTAGCAGACAAATAGACATTAATCTTCTATGTGTGAAA 1173
 QY 1081 AGAGTTTTTCAAGATATGCAATCTGAAATATTTGTTAGTCCAGCCATATATTTGCT 1140
 Db 1174 AGAGTTTTTCAAGATATGCAATCTGAAATATTTGTTAGTCCAGCCATATATTTGCT 1233
 QY 1141 GGTGGAGTATAGCTGACTATTTGAAATCAAAAGCAATTTTGGAAATTAATGTTTTC 1200
 Db 1234 GGTGGAGTATAGCTGACTATTTGAAATCAAAAGCAATTTTGGAAATTAATGTTTTC 1293
 QY 1201 ATATGCTGTTTCACTGTTATATGTTCCCTCAGAAATGCTGGAATTCCTTACTTATTTTA 1260
 Db 1294 ATATGCTGTTTCACTGTTATATGTTCCCTCAGAAATGCTGGAATTCCTTACTTATTTTA 1353
 QY 1261 CCTTATGCTATTTATAGCTTAAACATTAACCTGCTCCACATCAGACCTGTTTGTGAAA 1320
 Db 1354 CCTTATGCTATTTATAGCTTAAACATTAACCTGCTCCACATCAGACCTTGTGAAA 1413
 QY 1321 CTGAGTGTGTAAGCAATTTGTAATTTGCAATTTTACATCTTTCTGAAACAGACTTTT 1380
 Db 1414 CTGAGTGTGTAAGCAATTTGTAATTTGCAATTTTCTGAAACAGACTTTT 1473
 QY 1381 CAGTGGCCAAATAGTCCGACATTTCAAGGTTTATGTTGTAATTCAGTGTATTTTGA 1440
 Db 1474 CAGTGGCCAAATAGTCCGACATTTCAAGGTTTATGTTGTAATTCAGTGTATTTTGA 1533
 QY 1441 CTGTAAAAATGAGCTTATATA-TAGACATTTTCTACAAAGAACTGAAATAGAGGAAA 1499
 Db 1534 CTGTAAAAATGAGCTTATATA-TAGACATTTTCTACAAAGAACTGAAATAGAGGAAA 1593
 QY 1500 ACATGGAATTTCTTTTATGAGTGGTGGTCTTCAAAATTAATTAAGTTTATATA 1559
 Db 1594 ACATGGAATTTCTTTTATGAGTGGTGGTCTTCAAAATTAATTAAGTTTATATA 1653
 QY 1560 TATTTTAAACATATGTAAGAAATTAAGTGCAGAAAGAACTGGAAAGCTTAAGACTGCTT 1619
 Db 1654 TATTTTAAACATATGTAAGAAATTAAGTGCAGAAAGAACTGGAAAGCTTACTGCTT 1713
 QY 1620 CAANAAGCTGAATTAATGGAAGAAATTAANWMTTTCAGATATCTCAATGCGCCTNNKNA 1679
 Db 1714 CAANAAGCTGAATTAATGGAAGAAATTAANWMTTTCAGATATCTCAATGCGCCTCA 1773
 QY 1680 TGTGTGCCCTTNNCAANGCTTGGCAATGKTTNNWNTGNATAGTTNATTAANCTGGNN 1739

Db 1774 TGTGGCCCTTAAAGCTTGGGAATGTTTGTATATACAGATTATTAACCTGGGTA 1833
 QY 1740 TGTCTNNMATNA 1751
 Db 1834 TGTCTCAAAAAA 1845
 RESULT 3
 ABT05463
 ID ABT05463 standard; DNA; 1461 BP.
 XX
 AC ABR05463;
 XX
 DT 11-OCT-2002 (first entry)
 XX
 DE DNA of NOVX 9b SEQ ID No 21.
 XX
 KW Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipemic;
 KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
 KW tranquiliser; neuroleptic; antidiabetic; antilucer; antiinflammatory;
 KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
 KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
 KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
 KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
 KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
 KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
 KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
 KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
 KW gene therapy; gene; ds.
 XX
 OS Unidentified.
 XX
 PN WO200246409-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-US46586.
 XX
 PR 06-DEC-2000; 2000US-251660P.
 PR 12-DEC-2000; 2000US-255029P.
 PR 08-JAN-2001; 2001US-260326P.
 PR 24-JAN-2001; 2001US-263800P.
 PR 20-FEB-2001; 2001US-269942P.
 PR 24-APR-2001; 2001US-286183P.
 PR 20-AUG-2001; 2001US-313627P.
 PR 12-SEP-2001; 2001US-318712P.
 XX
 PA (CURA-) CURAGEN CORP.
 PA
 PI Guo X, Li L, Paturajan M, Shinkets RA, Casman SJ, Malyankar UM;
 PI Tcherev VT, Vernet CM, Spytek KA, Shenoy SG, Alsobrook JP;
 PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA;
 PI Boloid FL, Colman SD, Eisen AJ, Liu X, Padigaru M, Spederna SK;
 PI zerhusen BD;
 XX
 DR WPI: 2002-547774/58.
 DR P-PSDB; ABJ04650.
 XX
 PT Novel isolated polypeptide, designated NOVX, useful for treating or
 PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
 PT metabolic, neurodegenerative, immune and hematopoietic disorders -
 XX
 PS Claim 9; Page 84; 421pp; English.
 XX
 CC The invention relates to an isolated polypeptide, designated NOVX,
 CC comprising a sequence fully defined in the specification. The isolated
 CC protein, its encoding polynucleotide or an antibody created from the
 CC protein is useful in the manufacture of a medicament for treating a
 CC syndrome associated with a human disease, preferably a NOVX-associated
 CC disorder, or for treating or preventing a NOVX-associated disorder in a
 CC subject, preferably human. The isolated protein, its encoding

CC polynucleotide or an antibody created from the protein are also useful
CC for treating or preventing metabolic disorders, diabetes, obesity,
CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
CC disease, Parkinson's disorder, immune disorders, haematopoietic
CC disorders, and various dyslipidaemias, metabolic disturbances associated
CC with obesity, the metabolic syndrome X, wasting disorders associated with
CC chronic diseases, and cancer. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are useful for
CC treating or preventing neurological disorders such as epilepsy, stroke,
CC mental disorders including mood, anxiety, schizophrenic disorders,
CC disorders of vesicular transport such as cystic fibrosis, diabetes
CC mellitus, goiter, gastrointestinal disorders including ulcerative
CC colitis, other conditions associated with abnormal vesicle trafficking
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC arthritis. A cell comprising the vector of the invention is useful for
CC producing non-human transgenic animals. The polynucleotide of the
CC invention can be used to treat disorders by gene therapy. This
CC polynucleotide sequence represents the DNA encoding one of the isolated
CC NOVX proteins of the invention.

Sequence 1461 BP; 347 A; 298 C; 277 G; 539 T; 0 other;

Query Match 73.2%; Score 1359.8; DB 24; Length 1461;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1379; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 1 AAGGCGCAGCTAGAGGCTTACTGTTCTGCGCCGCTTGAAGCTTTTATGTC 60
DB 49 ATGGCCGACGTGAGAGTTACTATTTCTGCGCCGCTTGAAGCTTTTATGTC 108
QY 61 TGCCTCTCTCTCCGCTTCAAGCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 120
DB 109 TGCCTCTCTCTCCGCTTCAAGCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 168
QY 121 CACCTGCGCAGGCGGAGCGCTACTGAGGCGCATTTCTCCCTTCCAGTGGATCCC 180
DB 169 CACCTGCGCAGGCGGAGCGCTACTGAGGCGCATTTCTCCCTTCCAGTGGATCCC 228
QY 181 ATGATACTACATATCTGCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 229 ATGATACTACATATCTGCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288
QY 241 TGGATCTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 289 TGGATCTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348
QY 301 CTCTCTTTCAGTGGGCAACTCTTATTTACTATATTTGCTTCCAGAGGTACACC 360
DB 349 CTCTCTTTCAGTGGGCAACTCTTATTTACTATATTTGCTTCCAGAGGTACACC 408
QY 361 AGAACAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 409 AGAACAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
QY 421 CCAACACTTATTTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTT 480
DB 469 CCAACACTTATTTTATTTTATTTTATTTATTTATTTATTTATTTATTTT 528
QY 481 CTTTTCGATTTTGTGTGCTTATTTGGAATCATTTTACCTTCCCTGATTT 540
DB 529 CTTTTCGATTTTGTGTGCTTATTTGGAATCATTTTACCTTCCCTGATTT 588
QY 541 TGGGCTTATGTTTGGGCAACAATATCATCTGGGCTGCTTCTGTCAGGAGATGTC 600
DB 589 TGGGCTTATGTTTGGGCAACAATATCATCTGGGCTGCTTCTGTCAGGAGATGTC 648
QY 601 ATTGCAAAAAGTTAACTGAGGCTTGAAGAACTGAGCTCAAAAAGAGAGACAGACTT 660
DB 649 ATTGCAAAAAGTTAACTGAGGCTTGAAGAACTGAGCTCAAAAAGAGAGACAGACTT 708
QY 661 CCACTTATTTAAAGACCATTTGCAAGATTCAGAAAAATTTCTTATTTTGGCTTAT 720
DB 709 CCACTTATTTAAAGACCATTTGCAAGATTCAGAAAAATTTCTTATTTTGGCTTAT 768

QY 721 TCCATGCTCTTAAAAAAGTTGAGTATGCTTTCTGTTGACCTTGGCCCTACATCTTCTG 780
DB 769 TCCATGCTCTTAAAAAAGTTGAGTATGCTTTCTGTTGACCTTGGCCCTACATCTTCTG 828
QY 781 GGAATTCGTTTGTGCTTTTGTAGTAAATGAGTAAATGTTATTTGCGATCGAGT 840
DB 829 GGAATTCGTTTGTGCTTTTGTAGTAAATGAGTAAATGTTATTTGCGATCGAGT 888
QY 841 AGTATGAAGGCTGCTTCAATTTTCTCAATTTTCTTATTTTCTTATTTTCTTCTT 900
DB 889 AGTATGAAGGCTGCTTCAATTTTCTCAATTTTCTTATTTTCTTATTTTCTTCTT 948
QY 901 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
DB 949 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1008
QY 961 AAACATGAATTCGTTTGTGCTTGTAGTAAATGAGTAAATGTTATTTGGAATTC 1020
DB 1009 AAACATGAATTCGTTTGTGCTTGTAGTAAATGAGTAAATGTTATTTGGAATTC 1068
QY 1021 ACTTATGCTATTAATTAATCTGCTAGAGCAATATGATTAATCTTATGTCGGA 1080
DB 1069 ACTTATGCTATTAATTAATCTGCTAGAGCAATATGATTAATCTTATGTCGGA 1128
QY 1081 AGATTTTCAAGATATGCAATTTGAAATATTTGTTAGTTCAGGCTTATTTGCT 1140
DB 1129 AGATTTTCAAGATATGCAATTTGAAATATTTGTTAGTTCAGGCTTATTTGCT 1188
QY 1141 GGTGAGATATGCTGCTATTAATTAATGAAATGAAATGAAATGAAATGAAATG 1200
DB 1189 GGTGAGATATGCTGCTATTAATTAATGAAATGAAATGAAATGAAATGAAATG 1248
QY 1201 ATATGCTTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1260
DB 1249 ATATGCTTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1308
QY 1261 CCTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1320
DB 1309 CCTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1368
QY 1321 CTGAGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
DB 1369 CTGAGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1428
QY 1381 CAGTGGCAAAATGCTAGGACATTCAAAGT 1411
DB 1429 CAGTGGCAAAATGCTAGGACATTCAAAGT 1459

RESULT 4
ABT05462
ID ABT05462 standard; DNA; 1489 BP.
XX
AC ABT05462;
XX
DT 11-OCT-2002 (first entry)
XX
DE DNA of NOVX 9a SEQ ID NO 19.
XX
CYTOSTATIC; antidiabetic; anorectic; metabolic; nootropic; antilipemic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW neurotrophic; neuroleptic; antidiabetic; antitumor; antiinflammatory;
KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
KW gene therapy; gene; ds.

Db 781 TCCTCAGAACTGCTGGAATTCGTTACTTATTTTACCTTATGTCATTTATAGGCTTAA 840
 QY 1284 CATAACTCTGCTCCACATCCAGACTTGTGTGACACTGATGCTATGCAATTTGTTAA 1343
 Db 841 CATACCTCTGCTCCACATCCAGACTTGTGTGACACTGATGCTATGCAATTTGTTAA 900
 QY 1344 TTTTCATACCTTTTACATCTTTCTGACACAAAGCTTTTACAGGCGCAATTTAGTCAGACAT 1403
 Db 901 TTTTCATACCTTTTACATCTTTCTGACACAAAGCTTTTACAGGCGCAATTTAGTCAGACAT 960
 QY 1404 TCAAAGGTTTATGTCATATACGTATATTTTGAACGTGTAATAAGACTTATTA-T 1462
 Db 961 TCAAAGGTTTATGTCATATACGTATATTTTGAACGTGTAATAAGACTTATTA-T 1020
 QY 1463 AGACCATTTCTACAAAGAACACTGAATAGNGGAAACATGATTTCTTTTAGTGCA 1522
 Db 1021 AGACCATTTCTACAAAGAACACTGAATAGNGGAAACATGATTTCTTTTAGTGCA 1080
 QY 1523 GTGCGTCTTCAATTCATTTAGTTTATATATATTTTAAACATATGTAGAAT 1582
 Db 1081 GTGCGTCTTCAATTCATTTAGTTTATATATATTTTAAACATATGTAGAAT 1140
 QY 1583 TAAGTGGCAAGAACTGGGAAGCTTAGACCTGCTTCAAAAGCCTGAATATGGGAAA 1642
 Db 1141 TAAGTGGCAAGAACTGGGAAGCTTAGACCTGCTTCAAAAGCCTGAATATGGGAAA 1200
 QY 1643 TAAANMGTTTNCAGATATCTCATATCGCTCNCNNKATGNTGGCCCTTMCANAGCTTG 1702
 Db 1201 TAAATGTTTTCAGATATCTCATATCTCATATAGTTGGCCCTTAAAGCTTGG 1260
 QY 1703 GAATGKTNNMNTGNATAGTTTAACTGGGNNNTGCTNNMATNACT 1753
 Db 1261 GAATGTTTGTATGTACAAAGTTTATTAAGCTGGGTATGCTCATATTTACT 1311
 RESULT 6
 AAA87668
 ID AAA87668 standard: cDNA: 1165 BP.
 AC AAA87668;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE Human secreted protein gene 3 SEQ ID NO:13.
 XX
 KW Human; secreted protein; immunosuppressive; immunostimulant; neutrotrophic;
 KW antiinflammatory; cardiant; vulnerary; antilucer; anticonvulsant;
 KW antiparkinsonian; neuroprotective; antiviral; antibacterial; cytostatic;
 KW antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic;
 KW gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;
 KW cancer; immune system disorder; hyperproliferative disorder; infection;
 KW cardiovascular disorder; neurological disease; wound healing; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200043495-A2.
 XX
 PD 27-JUL-2000.
 XX
 PF 18-JAN-2000; 2000WO-US00903.
 XX
 PR 19-JAN-1999; 99US-0116330.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Moore PA;
 PI Komatsoulis G, Birse CE;
 XX
 DR WPI: 2000-499225/44.
 DR P-PSDB: AAB25667.
 XX
 PT New isolated polynucleotide encoding a secreted protein useful for
 preventing, treating or ameliorating a medical condition -

XX Claim 1; Page 380; 451pp; English.
 PS
 XX
 CC The polynucleotide sequences given in AAB87666 to AAA87708 encodes the
 CC human secreted proteins given in AAB25665 to AAB25755. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: immunosuppressive;
 CC immunostimulant; antiinflammatory; cardiant; vulnerary; antilucer;
 CC neutrotrophic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
 CC antibacterial; antiparasitic; thrombolytic; anticoagulant;
 CC antiarteriosclerotic and cytostatic. The secreted proteins and their
 CC polynucleotides can be used in gene therapy and as vaccines.
 CC chemotaxis-modulators and angiogenesis-modulators. The human secreted
 CC proteins and polynucleotides can be used for diagnosing (the
 CC susceptibility to) a pathological condition by determining the presence
 CC or absence of a mutation in the polynucleotide or determining the
 CC presence or amount of expression of the protein. The polynucleotides and
 CC proteins can also be used in the treatment and diagnosis of cancer,
 CC diseases of the immune system, hyperproliferative disorders,
 CC cardiovascular disorders and neurological disease. They can also be used
 CC to promote wound healing and to fight infection. AAA87657 to AAA87665 and
 CC AAB25664 represent sequences used in the exemplification of the present
 CC invention.
 SX
 SQ Sequence 1165 BP; 290 A; 262 C; 228 G; 385 T; 0 other;
 Query Match 48.4%; Score 899.6; DB 21; Length 1165;
 Best Local Similarity 97.4%; Pred. No. 2.4e-205;
 Matches 914; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
 QY 1 ATGGCGCAGCTAGAGGTTACTGTTCTCGGCGCCCTTGACCTGTACTTTTATGTCGC 60
 Db 136 ATGGCGCAGCTAGAGGTTACTGTTCTCGGCGCCCTTGACCTGTACTTTTATGTCGC 195
 QY 61 TGCCTCCTCTCCGCGCTTACGCGGCGGCGGAGAGCCCTACATGAGAGATCTTC 120
 Db 196 TGCCTCCTCTCTCCGCGCTTACGCGGCGGCGGAGAGCCCTACATGAGAGATCTTC 255
 QY 121 CACCTGCGTCAGGCGCAGCGCTACTGTGAGGCGCATTTCTCCCTTCCATGAGGATCC 180
 Db 256 CACCTGCGTCAGGCGCAGCGCTACTGTGAGGCGCATTTCTCCCTTCCATGAGGATCC 315
 QY 181 ATGATTTACTATTAACCTGCTGCTTACTGCTGTGACAGTGTGATTAACCTGCAATT 240
 Db 316 ATGATTTACTATTAACCTGCTGCTTACTGCTGTGACAGTGTGATTAACCTGCAATT 375
 QY 241 TGGATCTTTGATGTCGTCGAACATGTTGCTGCTCCATTTGGAGTCCAGATTTGTAT 300
 Db 376 TGGATCTTTGATGTCGTCGAACATGTTGCTGCTCCATTTGGAGTCCAGATTTGTAT 435
 QY 301 CTTCCTCTCAGTGTGGCACTTCTATTTACTATATTTGCTTTCCAAAGATCAACCC 360
 Db 436 CTTCCTCTCAGTGTGGCACTTCTATTTACTATATTTGCTTTCTGCAAGATCAACCC 495
 QY 361 AGAAACAAGGCTGCTCAAGTATCCAGAGACTTTGTCAACATTAACACTGACGATATT 420
 Db 496 AGAAACAAGGCTGCTCAAGTATCCAGAGACTTTGTCAACATTAACACTGACGATATT 555
 QY 421 CCAACACTTTATTTTATTTTAACTTCTTATATACAGAGAGATCATTTTAACT 480
 Db 556 CCAACACTTTATTTTATTTTAACTTCTTATATACAGAGAGATCATTTTAACT 615
 QY 481 CTTTTCGATATTTGATGTGCTTATATGAATCATAAACTTACGCTTCTGTGATTT 540
 Db 616 CTTTTCGATATTTGATGTGCTTATATGAATCATAAACTTACGCTTCTGTGATTT 675
 QY 541 TGTGGCTTCATGTTTGGCAAAATATCATCTGGGCTGCTTCTGTGCAAGGATGTC 600
 Db 676 TGTGGCTTCATGTTTGGCAAAATATCATCTGGGCTGCTTCTGTGCAAGGATGTC 735
 QY 601 ATTGCACAAAAGTATACGAGCTTGGAAAACCTAGCTACAAAGAAGAGAGACTT 660
 Db 736 ATTGCACAAAAGTATACGAGCTTGGAAAACCTAGCTACAAAGAAGAGAGACTT 795

OY 841 AGCATGAGCCTGCTTCATTTCCCTCACTATTTCTACTTTTTCATTACTCTCTT 900
 DB 968 AGTCATGAGCCTGCTTCATTTCCCTCACTATTTCTACTTTTTCATTACTCTCTT 1027
 OY 901 TTTTCCTTTCCCATCTCCCTGCTCCCTAGCAAAATTA 938
 DB 1028 TTTTCCTTTCCCATCTCCCTGCTCCACAAATTAATA 1065
 RESULT 8
 AAH06644
 ID AAH06644 standard; cDNA: 807 BP.
 AC AAH06644;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:3479.
 XX
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI, 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT full-length cDNAs -
 PS
 PS Claim 1; SEQ ID 3479; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 807 BP; 200 A; 140 C; 151 G; 312 T; 4 other;

Query Match 39.1%; Score 725.2; DB 22; Length 807;
 Best Local Similarity 95.4%; Pred. No. 1.2e-163;
 Matches 768; Conservative 0; Mismatches 34; Indels 3; Gaps 2;
 OY 444 CCTTATTTATACAGACGAGATCTATGTTTTTACTCTTTTGTGATATTTGATGCTCT 503
 DB 1 CCTTTATTTATACAGACGAGATCTATGTTTTTACTCTTTTGTGATATTTGATGCTCT 60
 OY 504 TTATGGAATCATTAACCTTACGCTTCCTTGATTTGTGGCTTCATGTTTGGCAAC 563
 DB 61 TTATGGAATCATTAACCTTACGCTTCCTTGATTTGTGGCTTCATGTTTGGCAAC 120
 OY 564 AAATATCATCTGGGCTGCTCTCTGTCAGGGAATGTCATTCACAAAAGTTAACTGAGC 623
 DB 121 AAATATCATCTGGGCTGCTCTCTGTCAGGGAATGTCATTCACAAAAGTTAACTGAGC 180
 OY 624 TTGGAACCTGAGCTACAAAAGAGAGACAGACTTCCACTTATTAAGACCTATTCG 683
 DB 181 TTGGAACCTGAGCTACAAAAGAGAGACAGACTTCCACTTATTAAGACCTATTCG 240
 OY 684 AGAATTCAGAAAATTTCTTCAGTTCTTTGGCTTATTCATGTCCTTTAAACCTTGAG 743
 DB 241 AGAATTCAGAAAATTTCTTCAGTTCTTTGGCTTATTCATGTCCTTTAAACCTTGAG 300
 OY 744 TATGCTTTTCTGTTTGACTTGCCCTACATCCTTCTGGGAATTTCTGTTTGTCTTTGT 803
 DB 301 TATGCTTTTCTGTTTGACTTGCCCTACATCCTTCTGGGAATTTCTGTTTGTCTTTGT 360
 OY 804 AGTAGTTAATGTTGGAATTTGTTATGGGATCGAGATGTCATGAAGCCTGCTTCATTT 863
 DB 361 AGTAGTTAATGTTGGAATTTGTTATGGGATCGAGATGTCATGAAGCCTGCTTCATTT 420
 OY 864 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTTCTTCTCAATCCCTGTC 923
 DB 421 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTTCTTCTCAATCCCTGTC 480
 OY 924 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTTCTTCTCAATCCCTGTC 983
 DB 481 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTTCTTCTCAATCCCTGTC 540
 OY 984 GGTACCTTACTGCTCTGCTTTTATAGTTGGAAATTCATGCTATTAATCTTCTCT 1043
 DB 541 GGTACCTTACTGCTCTGCTTTTATAGTTGGAAATTCATGCTATTAATCTTCTCT 600
 OY 1044 AGCAGACATATGACATTTATCTTCTATGTTGGAAAGATTTTCAAGATATGCAAT 1103
 DB 601 AGCAGACATATGACATTTATCTTCTATGTTGGAAAGATTTTCAAGATATGCAAC 660
 OY 1104 TCTGAAATATTTGTTAGTTCCAGCCTATATATTTGCTGGTGGAGATATGCTACAT 1163
 DB 661 TGT-AAATATTTGGATGTTCCAGCCTATATATTTGCTGGTGGAGATATGCTACAT 719
 OY 1164 GAATTCAGAACCAATTTTGTGAATTTATGTTTTCATATGCTT--GTCATGTTATA 1221
 DB 720 GAATTCAGAACCAATTTTGTGAATTTATGTTTTCATATGCTT--GTCATGTTATA 779
 OY 1222 GTTCTCAGAAACTGCTGGAATTTTC 1246
 DB 780 GTTCTCAGAAACTGCTGGAATTTTC 804
 RESULT 9
 AB211754
 ID AB211754 standard; cDNA: 843 BP.
 AC AB211754;
 XX
 DT 20-JAN-2003 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 636.
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;

KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW hemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 KW antirheumatic; gene; ss.
 XX Homo sapiens.
 OS
 PN NO200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002MO-US05095.
 XX
 PR 05-MAR-2001; 2001US-0799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Dimanac RT;
 DR WPI: 2002-759812/82.
 DR P-PSDB; ABP69537.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders -
 XX
 PS Claim 1; SEQ ID NO 636; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (AB211119-AB212066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 843 BP; 178 A; 209 C; 191 G; 265 T; 0 other;
 Query Match 27.4%; Score 509.4; DB 24; Length 843;
 Best Local Similarity 95.6%; Pred. No. 5,7e-112;
 Matches 567; Conservative 0; Mismatches 21; Indels 5; Gaps 4;
 OY 1 ATGGCCGAGCTAAGGGTACTGTTCTGCGCGCCCTTAGCGTACCTTTTAACTGTC 60
 DB 252 ATGGCGAGCTAAGGGTACTGTTCTGCGCGCCCTTAGCGTACCTTTTAACTGTC 311
 OY 61 TGGCTCTCTTTCGCGCTTCAGCGGCGCTCGAGAGCCATACATGAGCGAGATCTTC 120
 DB 312 TGGCTCTCTTTCGCGCTTCAGCGGCGCTCGAGAGCCATACATGAGCGAGATCTTC 371
 OY 121 CACCTGCTTCAGCGAGCGCTACTGTGAGGCGCATTTTCCTTTCCAGTGGATCCC 180
 DB 372 CACCTGCTTCAGCGAGCGCTACTGTGAGGCGCATTTTCCTTTCCAGTGGATCCC 431
 OY 181 ATGATTAACATATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB 432 ATGATTAACATATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491

OY 241 TGGATCTTTGGATGCTCTGAACAGTGTGCTGCGCCATTGGAGTCTGATTTGTTAA 300
 DB 492 TGGATCTTTGGATGCTCTGAACAGTGTGCTGCGCCATTGGAGTCTGATTTGTTAA 551
 OY 301 CTCTCTTCAGCTGTTGGCAACTCTATATTACTATATTGCTTTCCAAAGATACCAACC 360
 DB 552 CTCTCTTCAGCTGTTGGCAACTCTATATTACTATATTGCTTTCCAAAGATACCAACC 611
 OY 361 AGAACAAGCTGCTCTAATATTCAGAGAGTCTTGCAACATTACACTAG-CAGTATT 419
 DB 612 AGAACAAGCTGCTCTAATATTCAGAGAGTCTTGCAACATTACAGAGAGTATT 671
 OY 420 TCCACACTTATTTTATTTTACCTCTTATTTATACAGAGAGAGATCTATGTTTATAC 479
 DB 672 TCCACACTTATTTTATTTTACCTCTTATTTATACAGAGAGAGATCTATGTTTATAC 731
 OY 480 TCTTTTGCATATTGATGTGCTTTATGAGAAATCATMAAATTCACGCTTCCTTGAT 539
 DB 732 TCTTTTGCATATTGATGTGCTTTATG- AATCATMAAATTCACGCTTCCTTGAT 790
 OY 540 TTGTGGCTTATGTTTGGCAACAATATATCATCTGGGCTGCTTCTGTCGAG 592
 DB 791 TTGT-GCTTCATGTTCCGG--AAACAACAATCATCTGGGCTGCTTCTGTCGAG 840

RESULT 10
 AA258638
 ID AA258638 standard; DNA; 539 BP.
 XX
 AC AA258638;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE Pancreatic cancer and dysplasia up-regulated polynucleotide #5.
 XX
 KW Pancreatic cancer; dysplasia; forensic; genetic analysis; cancer; bone;
 KW brain tumour; breast cancer; endocrine system cancer; gastrointestinal;
 KW genitourinary; gynecological; leukemia; Hodgkin's lymphoma; pediatric;
 KW skin; urinary tract; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO967386-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 22-JUN-1999; 99MO-US14036.
 XX
 PR 23-JUN-1998; 98US-0090391.
 PR 03-FEB-1999; 99US-0118570.
 PR 21-JUN-1999; 99US-0337171.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Kennedy GC;
 XX
 PT WPI: 2000-147209/13.
 XX
 PS Claim 1; Page 38-39; 43pp; English.
 XX
 CC The invention provides polynucleotide sequences (AA258634-648) that are
 CC differentially expressed in pancreatic cancer and dysplasia. The
 CC polynucleotides and polypeptides encoded by them are useful for the
 CC diagnosis and treatment of pancreatic cancer and dysplasia. The
 CC polynucleotides are also useful in forensics, genetic analysis, mapping,
 CC and diagnostic applications if the corresponding region of a gene is
 CC polymorphic in the human population. Therapeutic compositions comprising
 CC antibodies specific for the polypeptides are useful for treating
 CC pancreatic cancer and pancreatic dysplasia as well as other types of

CC cancer e.g. bone cancer, brain tumours, breast cancer, endocrine system
CC cancers e.g. thyroid, pituitary and adrenal glands and the pancreatic
CC islets, gastrointestinal cancers, genitourinary cancers, gynecological
CC cancers, head and neck cancers, leukemia, lymphomas including Hodgkin's
CC and non-Hodgkin's lymphoma, metastatic cancer, myelomas, sarcomas, skin
CC cancer, urinary tract cancers and pediatric cancers. The present
CC sequence represents a polynucleotide sequence that is up-regulated in
CC both dysplasia and pancreatic cancer.
XX
SQ Sequence 539 BP; 165 A; 79 C; 90 G; 205 T; 0 other:

Query Match 27.0%; Score 502; DB 21; Length 539;
Best Local Similarity 96.8%; Pred. No. 2.8e-110;
Matches 522; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 1051 AATAGACATTTACTTCTATGTGGAAGAGTTTCAAGATGCAATTCGTGAA 1110
DB 1 AATAGACATTTACTTCTATGTGGAAGAGTTTCAAGATGCAATTCGTGAA 60
QY 1111 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGAGTATAGCTCATTTGAAATCA 1170
DB 61 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGAGTATAGCTCATTTGAAATCA 120
QY 1171 AAGCAATTTTGGAAATTAATGTTTTCATATGCTGTCATGTATAGTTCCTGAG 1230
DB 121 AAGCAATTTTGGAAATTAATGTTTTCATATGCTGTCATGTATAGTTCCTGAG 180
QY 1231 AAACGTGCAATTTGCTTACTTCTATTTTACCTTATGCTATATAGCTTAACATACT 1290
DB 181 AAACGTGCAATTTGCTTACTTCTATTTTACCTTATGCTATATAGCTTAACATACT 240
QY 1291 CTGCTCCCAATCCAGACTGTTTGTGTAAGTGAAGTCTATGCAATTTGTAATTCATA 1350
DB 241 CTGCTCCCAATCCAGACTGTTTGTGTAAGTGAAGTCTATGCAATTTGTAATTCATA 300
QY 1351 ACTTTTTCATCTTTCGAAAGAAAGCTTTGAGTGGCCAAATGACAGACATTCAAAGG 1410
DB 301 ACTTTTTCATCTTTCGAAAGAAAGCTTTGAGTGGCCAAATGACAGACATTCAAAGG 360
QY 1411 TTTATGTTGTAATATAGTATATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1469
DB 361 TTTATGTTGTAATATAGTATATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 420
QY 1470 TTTCAACAAGAAACAATGAAAGGAAAGAAAGAAAGAAAGAAAGTCTTAAAGTGG 1529
DB 421 TTTCAACAAGAAACAATGAAAGGAAAGAAAGAAAGAAAGTCTTAAAGTGG 480
QY 1530 TCTTCAATTTACATTTAGTTTCTTAAATATATTTTAAACATATGTAAGATTAAGTG 1588
DB 481 TCTTCAATTTACATTTAGTTTCTTAAATATATTTTAAACATATGTAAGATTAAGTG 539

RESULT 11
AAK89423
ID AAK89423 standard; DNA; 387 BP.
XX
AC AAK89423;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 2999.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.

XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0188874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225265.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.

PT	Poly nucleotides encoding digestive system antigens, useful for diagnosing, treating, preventing and/or prognosing disorders of the
XX	
DR	WPI; 2001-502630/55.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Rosen CA, Barash SC, Ruben SM;
XX	
XX	
PR	02-OCT-2000; 2000US-0236802.
PR	02-OCT-2000; 2000US-0237037.
PR	02-OCT-2000; 2000US-0237038.
PR	02-OCT-2000; 2000US-0237039.
PR	02-OCT-2000; 2000US-0237040.
PR	13-OCT-2000; 2000US-0239935.
PR	13-OCT-2000; 2000US-0239937.
PR	20-OCT-2000; 2000US-0240960.
PR	20-OCT-2000; 2000US-0241221.
PR	20-OCT-2000; 2000US-0241785.
PR	20-OCT-2000; 2000US-0241786.
PR	20-OCT-2000; 2000US-0241787.
PR	20-OCT-2000; 2000US-0241808.
PR	20-OCT-2000; 2000US-0241809.
PR	20-OCT-2000; 2000US-0241826.
PR	01-NOV-2000; 2000US-0244617.
PR	08-NOV-2000; 2000US-0246474.
PR	08-NOV-2000; 2000US-0246475.
PR	08-NOV-2000; 2000US-0246476.
PR	08-NOV-2000; 2000US-0246477.
PR	08-NOV-2000; 2000US-0246478.
PR	08-NOV-2000; 2000US-0246523.
PR	08-NOV-2000; 2000US-0246524.
PR	08-NOV-2000; 2000US-0246525.
PR	08-NOV-2000; 2000US-0246526.
PR	08-NOV-2000; 2000US-0246527.
PR	08-NOV-2000; 2000US-0246528.
PR	08-NOV-2000; 2000US-0246532.
PR	08-NOV-2000; 2000US-0246609.
PR	08-NOV-2000; 2000US-0246610.
PR	08-NOV-2000; 2000US-0246611.
PR	08-NOV-2000; 2000US-0246613.
PR	17-NOV-2000; 2000US-0249207.
PR	17-NOV-2000; 2000US-0249208.
PR	17-NOV-2000; 2000US-0249209.
PR	17-NOV-2000; 2000US-0249210.
PR	17-NOV-2000; 2000US-0249211.
PR	17-NOV-2000; 2000US-0249212.
PR	17-NOV-2000; 2000US-0249213.
PR	17-NOV-2000; 2000US-0249214.
PR	17-NOV-2000; 2000US-0249215.
PR	17-NOV-2000; 2000US-0249216.
PR	17-NOV-2000; 2000US-0249217.
PR	17-NOV-2000; 2000US-0249218.
PR	17-NOV-2000; 2000US-0249244.
PR	17-NOV-2000; 2000US-0249245.
PR	17-NOV-2000; 2000US-0249264.
PR	17-NOV-2000; 2000US-0249265.
PR	17-NOV-2000; 2000US-0249297.
PR	17-NOV-2000; 2000US-0249299.
PR	17-NOV-2000; 2000US-0249310.
PR	01-DEC-2000; 2000US-0250160.
PR	01-DEC-2000; 2000US-0250391.
PR	05-DEC-2000; 2000US-0251030.
PR	05-DEC-2000; 2000US-0251988.
PR	05-DEC-2000; 2000US-0256719.
PR	06-DEC-2000; 2000US-0251479.
PR	08-DEC-2000; 2000US-0251856.
PR	08-DEC-2000; 2000US-0251868.
PR	08-DEC-2000; 2000US-0251869.
PR	08-DEC-2000; 2000US-0251989.
PR	08-DEC-2000; 2000US-0251990.
PR	11-DEC-2000; 2000US-0254097.
PR	05-JAN-2001; 2001US-0259678.
XX	

PT	digestive system, particularly cancer and cancer metastases -
PS	Disclosure; SEQ ID NO 2999; 986bp; English.
XX	
CC	The present invention provides the protein and coding sequences of a
CC	number of human digestive system antigens. These can be used in the
CC	diagnosis, treatment and prevention of digestive system disorders,
CC	including cancer, Meckel's diverticulum, bacterial or parasitic
CC	infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC	ulcerative colitis. The present sequence is a genomic DNA fragment
CC	encoding a digestive system antigen of the invention.
XX	
SQ	Sequence 387 BP; 133 A; 60 C; 65 G; 129 T; 0 other;
Query Match	18.9%; Score 350.4; DB 22; Length 387;
Best Local Similarity	95.1%; Pred. No. 4.9e-74;
Matches	368; Conservative 1; Mismatches 17; Indels 1; Gaps 1;
OY	1280 TTTAACATTAAGCTCTGCTCCACATCCAGACTGTTTGTGAACGTAGTTCATCAATTG 13399
DB	1 TTTAACATTAAGCTCTGCTCCACATCCAGACTGTTTGTGAACGTAGTTCATCAATTG 60
OY	1340 TTTAATTTCAATCACTTTTACATCTTTTCGAGACAGACTTTTCATGTGGCCAAATAGTCAG 13999
DB	61 TTTAATTTCAATCACTTTTTCATCTTTTCGAGACAGACTTTTCATGTGGCCAAATAGTCAG 120
OY	1400 ACATTTCAAAGSTTATGTGTAATATCAGTATTTTGAACSTGTAAATAATGACTTAAT 14599
DB	121 ACATTTCAAAGSTTATGTGTAATATCAGTATTTTGAACSTGTGAAAAATGACTTAAT 180
OY	1460 AA-TAGACCATTTCTACAAAGAACACTGAATAGGNGGAAAAACATGAAATTCCTTTAGG 1518
DB	181 AATTAGACCATTTCTACAAAGAACACTGAATAGTGGAAAAACATGAAATTCCTTTAGG 240
OY	1519 TGCAGTGGTGGTCTTCACAAATACATTAAGTTTTTTAATATATATTTTAACATATGTAA 1578
DB	241 TGCAGTGGTGGTCTTCACAAATACATTAAGTTTTTTAATATATATTTTAACATATGTAA 300
OY	1579 AAATTAAGTGCACAAAGAACTGGGAAAGCTTAAGACCTGCTTCAANGCCTGAATATAGG 1638
DB	301 AAATTAAGTGCACAAAGAACTGAGAAAGCTTAAGACCTGCTTCAANGCCTGAATATAGG 360
OY	1639 AAAATTAAMNMGTTTTNCAGATATCTCA 1665
DB	361 AAAATTAAMNMGTTTTTCAGATATCTCA 387
RESULT 12	
ID	AAF93488
AC	AAF93488 standard; cDNA; 469 BP.
XX	
DT	21-MAY-2001 (first entry)
DE	CDNA encoding SRT protein isolated from adenocarcinoma tissue SEQ ID 309.
XX	
XX	Human; SRT; gene therapy; gene mapping; tissue typing; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200107611-A2.
PD	01-FEB-2001.
XX	
PF	21-JUL-2000; 2000WO-US20006.
XX	
PR	26-JUL-1999; 99US-0145701.
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Baker KP, Goddard A, Wood WI;
XX	

DR WPI: 2001-112729/12.
XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful
PT for production of recombinant SRT polypeptides, gene mapping,
PR diagnosing genetic disorders and for gene therapy -
XX
XX Claim 2; Fig 309; 663pp: English.
XX
CC Sequences AAF93180 - AAF93143 represent polynucleotide sequences encoding
CC human SRT proteins. The cDNA sequences are isolated from various
CC different human tissue cDNA libraries. The invention relates to a method
CC for detecting cDNA encoding an SRT protein, a vector containing cDNA
CC encoding SRT, a host cell transformed with the vector, an isolated SRT
CC polypeptide, and an antibody which binds to SRT. The polynucleotide
CC sequence can be used in gene therapy and is useful in the recombinant
CC production of SRT polypeptides, as a hybridisation probe to screen
CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to
CC map the gene encoding the SRT polypeptides and analysing genetic
CC disorders, tissue typing and disease tissue detection. The SRT
CC polynucleotide sequences can be used in polymerase chain reaction,
CC screening for new therapeutic molecules and generation of antisense RNA
CC and DNA.
XX
XX Sequence 469 BP; 76 A; 125 C; 119 G; 146 T; 3 other;
SO
Query Match 18.5%; Score 343.2; DB 22; Length 469;
Best Local Similarity 98.6%; Pred. No. 2.8e-72;
Matches 345; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 ATGGCGCAGCTAGAGGCTTACTGTTTCGCGCCGCTGAGCTGACTGTTTACTGCTCC 60
DB 120 ATGGCGCAGCTAGAGGCTTACTGTTTCGCGCCGCTGAGCTGACTGTTTACTGCTCC 179
OY 61 TGCCCTCCTCTCTCCGCTTACGCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 120
DB 180 TGCCCTCCTCTCTCCGCTTACGCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 239
OY 121 CACCTGCGCTAGCGCAGCGCTACTGTGAGGGCCATTTCCTTTCCAGTGGATCC 180
DB 240 CACCTGCGCTAGCGCAGCGCTACTGTGAGGGCCATTTCCTTTCCAGTGGATCC 299
OY 181 ATGATTTACTACATACCTGCTTGTACCTGTCAGTGTGAGAGGTCGTAACCTGCCATT 240
DB 300 ATGATTTACTACATACCTGCTTGTACCTGTCAGTGTGAGAGGTCGTAACCTGCCATT 359
OY 241 TGCATCTTTGGATGTCGACATGTTGTCGTCATTTGGAGTGCATGTTGTTAT 300
DB 360 TGCATCTTTGGATGTCGACATGTTGTCGTCATTTGGAGTGCATGTTGTTAT 419
OY 301 CTTCCTCTCAGTGTGGCACTTCTATTACTATATTTGCTTTCCACAA 350
DB 420 CTTCCTCTCAGTGTGGCACTTCTATTACTATATTTGCTTTCCACAA 469
RESULT 13
AAF93569
ID AAF93569 standard; cDNA: 442 BP.
AC AAF93569;
XX
XX 21-MAY-2001 (first entry)
XX
DE Lung carcinoma cDNA encoding SRT protein SEQ ID 390.
XX
XX Human; SRT; gene therapy; gene mapping; tissue typing; ss.
XX
XX Homo sapiens.
XX
XX WO200107611-A2.
XX
XX 01-FEB-2001.
XX
XX 21-JUL-2000; 2000MO-US20006.
PF

XX
PR 26-JUL-1999; 99US-0145701.
XX
XX (GENE) GENENTECH INC.
PA
XX Baker KP, Goddard A, Wood WI;
PI
XX WPI: 2001-112729/12.
XX
XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful
PT for production of recombinant SRT polypeptides, gene mapping,
PR diagnosing genetic disorders and for gene therapy -
XX
XX Claim 2; Fig 390; 663pp: English.
XX
CC Sequences AAF93180 - AAF93143 represent polynucleotide sequences encoding
CC human SRT proteins. The cDNA sequences are isolated from various
CC different human tissue cDNA libraries. The invention relates to a method
CC for detecting cDNA encoding an SRT protein, a vector containing cDNA
CC encoding SRT, a host cell transformed with the vector, an isolated SRT
CC polypeptide, and an antibody which binds to SRT. The polynucleotide
CC sequence can be used in gene therapy and is useful in the recombinant
CC production of SRT polypeptides, as a hybridisation probe to screen
CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to
CC map the gene encoding the SRT polypeptides and analysing genetic
CC disorders, tissue typing and disease tissue detection. The SRT
CC polynucleotide sequences can be used in polymerase chain reaction,
CC screening for new therapeutic molecules and generation of antisense RNA
CC and DNA.
XX
XX Sequence 442 BP; 73 A; 112 C; 115 G; 140 T; 2 other;
SO
Query Match 18.1%; Score 336.4; DB 22; Length 442;
Best Local Similarity 99.1%; Pred. No. 1.2e-70;
Matches 337; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 11 TAGAGGCTTACTGTTTCTCGCGCGCTGAGCTGACTGTTTACTGCTCTCTCT 70
DB 103 TAGAGGCTTACTGTTTCTCGCGCGCTGAGCTGACTGTTTACTGCTCTCTCTCT 162
OY 71 TCTCCGCTTACGCGCGCGCTGCGAGAGCCCTACATGAGAGATGTCACCTGCCCTC 130
DB 163 TCTCCGCTTACGCGCGCGCTGCGAGAGCCCTACATGAGAGATGTCACCTGCCCTC 222
OY 131 AGGCGCAGCGCTACTGTGAGGGCCATTTCCTTTCCAGTGGAGTCCATGATTA 190
DB 223 AGGCGCAGCGCTACTGTGAGGGCCATTTCCTTTCCAGTGGAGTCCATGATTA 282
OY 191 CATTAAGCTGGCTTACTGCTGTGTCAGTGTGAGAGTGTCAACCTGCCATTGGATCTTG 250
DB 283 CATTAAGCTGGCTTACTGCTGTGTCAGTGTGAGAGTGTCAACCTGCCATTGGATCTTG 342
OY 251 GATGATCTGACATGTTGTCGTCATTTGGAGTGCATGTTGTTAATCTCTCTCA 310
DB 343 GATGATCTGACATGTTGTCGTCATTTGGAGTGCATGTTGTTAATCTCTCTCA 402
OY 311 GTGTTGGCACTTCTATTACTATATTTGCTTTCCACAA 350
DB 403 GTGTTGGCACTTCTATTACTATATTTGCTTTCCACAA 442
RESULT 14
AAH98971
ID AAH98971 standard; cDNA: 348 BP.
AC AAH98971;
XX
XX 12-OCT-2001 (first entry)
XX
XX Rat EST-derived coding sequence SEQ ID NO: 828.
XX
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KH

KM diagnostics; forensic test; gene mapping; genetic disorder;
XX biodiversity; gene therapy; nutrition; ss.
XX Rattus norvegicus.
XX WO200154477-A2.
XX 02-AUG-2001.
XX 25-JAN-2001; 2001MO-US02687.
XX 25-JAN-2000; 2000US-0491404.
XX 17-JUL-2000; 2000US-0617746.
XX 03-AUG-2000; 2000US-0631451.
XX 15-SEP-2000; 2000US-0663870.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX WPI: 2001-476164/51.
XX P-PSDB; AAM24312.
XX DR P-PSDB; AAM24312.
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use -
XX Claim 1; Page 691; 1275pp; English.
XX The present invention provides the protein and coding sequences of novel
XX proteins from a variety of organisms, including human, dog, cat, horse,
XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX from the organism of interest. They can be used in diagnostics,
XX forensics, gene mapping, identification of mutations, to assess
XX biodiversity and for nutritional purposes. The present sequence is a cDNA
XX of the invention.
XX Sequence 348 BP; 69 A; 65 C; 58 G; 156 T; 0 other;
XX
XX Query Match 18.0%; Score 334.2; DB 22; Length 348;
XX Best Local Similarity 97.7%; Pred. No. 3.5e-70;
XX Matches 339; Conservative 0; Mismatches 8; Indels 0; Caps 0;
XX
XX 670 AAAGGACATTCGACAGATTCAGAAAATCTCAGTTCTTTGGTTATTCAGATCC 729
XX 1 AAAGGACATTCGACAGATTCAGAAAATCTCAGTTCTTTGGTTATTCAGATCC 60
XX
XX 730 TTTTAAACCTTGAGTATGCTTTCTGTTGACTTGGCCCTACATCCCTCGGATTTCTG 789
XX 61 TTTTAAACCTTGAGTATGCTTTCTGTTGACTTGGCCCTACATCCCTCGGATTTCTG 120
XX
XX 790 TTTTGGCTTTTGTAGTAAATGCGAATGTTATTTGGCATCGAGTACATGAA 849
XX 121 TTTTGGCTTTTGTAGTAAATGCGAATGTTATTTGGCATCGAGTACATGAA 180
XX
XX 850 GCGTGCCTCATTTCTCTCAACATTTCTACTTTTTTTCATTTCTCTTTTCTTTTCT 909
XX 181 GCGTGCCTCATTTCTCTCAACATTTCTACTTTTTTTCATTTCTCTTTTCTTTTCT 240
XX
XX 910 CCTCATCTCCTGTCCTCTAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCT 969
XX 241 CCTCATCTCCTGTCCTCTAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCT 300
XX
XX 970 ATTCTGTTTGGTGGTGTACCTTAAGTCTGCTGTTTCTTCTTCTTCTTCTTCTTCT 1016
XX 301 ATTCTGTTTGGTGGTGTACCTTAAGTCTGCTGTTTCTTCTTCTTCTTCTTCTTCT 347
XX
XX RESULT 15
XX AAV87898
XX ID AAV87898 standard; cDNA; 283 BP.

AC AAV87898;
XX
XX 12-FEB-1999 (first entry)
XX
XX EST clone FE311.
XX
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
XX Homo sapiens.
XX WO9845437-A2.
XX
XX 15-OCT-1998.
XX
XX 10-APR-1998; 98WO-US06956.
XX
XX 10-APR-1997; 97US-0837312.
XX
XX (GENM) GENETICS INST INC.
XX
XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racine LA, Spaulding V, Treacy M;
XX WPI: 1999-070078/06.
XX
XX New polynucleotides encoding human secreted proteins - derived from
XX e.g. human blood, kidney, foetal lung, placenta, testes, brain,
XX ovary, pituitary, retina and colon cDNA libraries
XX
XX Claim 1; Page 216; 641pp; English.
XX
XX The present sequence represents an expressed sequence tag (EST), and is
XX a polynucleotide of the invention. The polynucleotides of the invention
XX are all secreted EST sequences isolated from a variety of human tissue
XX sources. The EST sequences and proteins encoded by them are predicted to
XX have useful biological activities which would make them suitable for
XX treating, preventing or ameliorating medical conditions in humans and
XX animals, although no supporting data is given. Suggested activities
XX include nutritional activity, immune stimulating or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition
XX activity. The EST sequences are also stated to be useful for gene
XX therapy.
XX
XX Sequence 283 BP; 75 A; 54 C; 54 G; 100 T; 0 other;
XX
XX Query Match 13.9%; Score 258.4; DB 20; Length 283;
XX Best Local Similarity 97.8%; Pred. No. 4.6e-52;
XX Matches 262; Conservative 0; Mismatches 6; Indels 0; Caps 0;
XX
XX 371 CTGCTCAAGTATTCACAGAGTCTGTGCAATTAACAGTATTCACAGACTTT 430
XX 12 CTGCTCAAGTATTCACAGAGTCTGTGCAATTAACAGTATTCACAGACTTT 71
XX
XX 431 ATTTTAACTTCCTTTATTAACAGAGCAGATCTATGTTTCTTCTTCTTCTTCTTCT 490
XX 72 ATTTTAACTTCCTTTATTAACAGAGCAGATCTATGTTTCTTCTTCTTCTTCTTCT 131
XX
XX 491 ATTTGATGTGTCTTTATGGAATCATTAACCTTCAGCTTCTTGGATTTGGCTTCA 550
XX 132 ATTTGATGTGTCTTTATGGAATCATTAACCTTCAGCTTCTTGGATTTGGCTTCA 191
XX
XX 551 TGTTCGGCAAAATATCATCTGGGCTGCTCTGTCAGGAATGTCATTGCACAAA 610
XX 192 TGTTCGGCAAAATATCATCTGGGCTGCTCTGTCAGGAATGTCATTGCACAAA 251
XX
XX 611 AGTTAACTGAGGCTTGGAAAACGTGAGCT 638
XX
XX

Db 252 AGTTACGGAGGCTTGGGAAC TGAGCT 279

Search completed: September 23, 2003, 11:15:36
Job time : 627 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 4775 Seconds
(without alignments)
9452.023 Million cell updates/sec

Title: US-10-000-151B-1

Perfect score: 1857

Sequence: 1 atggcgcagctagtaggggtta.....aagtaaatatgncncnaaa 1857

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estbda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1574	84.8	1833	11	BC022336 Homo sapi
2	872.4	47.0	974	9	AL578231 AL578231
3	760	40.9	958	9	AL555819 AL555819
4	639	34.4	784	12	BI770281 BI770281

5	634.2	34.2	821	12	BI825977
6	618.6	32.3	911	14	CD252003
c	596	32.1	1019	10	BE738701
8	580.4	31.3	1201	9	AL519606
9	527.6	28.4	667	14	CB483386
10	509.4	27.4	667	10	BG613309
c	503.4	27.1	694	12	BI768769
11	469.8	25.3	560	28	AO599053
c	464	25.0	1126	13	BX401726
13	461	24.8	530	28	AO512968
14	459	24.7	724	14	BY740968
15	431	23.2	527	2	HS006877
c	425	22.9	483	28	AO143712
17	425	22.8	629	14	BY722936
18	424	22.8	629	14	BY722936
19	423.8	22.8	531	12	BG835696
20	421.6	22.7	659	14	BY734976
21	413.4	22.3	883	13	BG133687
22	401	21.6	535	10	BG692134
c	398.8	21.5	640	10	BB866047
23	391	21.1	468	14	CB158813
c	379.6	20.4	775	9	AJ451344
25	377	20.3	783	13	B0113244
26	374	20.1	816	9	AJ445978
27	368.4	19.8	571	9	AL712776
28	365.8	19.7	862	10	BE738133
29	365.8	19.7	862	10	BE738133
30	361.2	19.5	709	13	B0305035
31	360.8	19.4	717	14	BY748369
32	356.8	19.2	491	12	BM149103
33	349	18.8	507	12	BM194113
c	337.8	18.2	721	10	BF031923
34	303	16.3	349	10	BF197055
c	303	16.3	720	13	B0313974
35	296.6	16.0	702	9	AL646668
37	296.6	16.0	623	14	BY745545
38	296.4	16.0	649	14	BY749182
39	296.4	15.9	454	14	CB784843
c	294.8	15.7	500	28	AO118630
40	291.8	15.7	443	12	BM146779
c	290.8	15.7	443	12	BM146779
42	272.4	14.7	889	13	B0912625
43	272.4	14.7	889	13	B0912625
44	269	14.5	428	9	AA223584
45	268.6	14.5	290	9	AW297936

ALIGNMENTS

RESULT 1
BC022336
LOCUS
DEFINITION Homo sapiens, hypothetical protein FLJ14751, clone IMAGE:4771940,
ACCESSION BC022336
VERSION BC022336.1
KEYWORDS GI:18490846
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 1833)
TITLE Strausberg, R.
Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov

COMMENT
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Sequencing group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdedpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMIL at: <http://image.llnl.gov>
 Series: IRLM Plate: 37 Row: m Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14249543
 This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers

1..1833
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /db_xref="IMAGE:4771940"
 /clone="IMAGE:4771940"
 /issue_type="Testis, embryonal carcinoma"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"

BASE COUNT 480 a 361 c 344 g 648 t
 ORIGIN

Query Match 84.8%; Score 1574; DB 11; Length 1833;
 Best Local Similarity 97.1%; Pred. No. 5e-237;
 Matches 1620; Conservative 1; Mismatches 46; Indels 2; Gaps 2;

```

OY 1 ATGGGGGAGCTAGAGGTTACTGTTTCGCGCGCTGAGCTGACCTTTTATGCTCC 60
DB 140 ATGGGCGAGCTGGAAGGTTACTATTCTCGCGCGCTGAGCTGACCTTTTATGATCC 199
OY 61 TGCCTCTCTCTCCGCTTCAGCGCGGCTGCGAGAGCCCTACATGAGCAGATCTTC 120
DB 200 TGCCTCTCTCTCCGCTTCAGCGCGGCTGCGAGAGCCCTACATGAGCAGATCTTC 259
OY 121 CACCTGCTCAGCGCGAGGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGAGTCC 180
DB 260 CACCTGCTCAGCGCGAGGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGAGTCC 319
OY 181 ATGATTACTACATTACCTGGCTGTACCGGTGTACGTTGAGTGGTCAACCTCCANT 240
DB 320 ATGATTACTACATTACCTGGCTGTACCGGTGTACGTTGAGTGGTCAACCTCCANT 379
OY 241 TGGATCTTGGATGCTGGAACATGTTGTCTGCTCCATTTGGATGCTGATTTGTAAT 300
DB 380 TGGATCTTGGATGCTGGAACATGTTGTCTGCTCCATTTGGATGCTGATTTGTAAT 439
OY 301 CTTCCTCTCAGTTGGGCACTTCTATTATTTGCTTTTCCACAGGTACAAACC 360
DB 440 CTTCCTCTCAGTTGGGCACTTCTATTATTTGCTTTTCCACAGGTACAAACC 499
OY 361 AGAACAAGGCGCTCAAGTATCCAGAGCTGTGCAACATTAACACTACAGATATT 420
DB 500 AGAACAAGGCGCTCAAGTATCCAGAGCTGTGCAACATTAACACTACAGATATT 559
OY 421 CCACACTTTATTTTAACTCTTATTTATACAGAGCAGATCTTATTTTACT 480
DB 560 CCACACTTTATTTTAACTCTTATTTATACAGAGCAGATCTTATTTTACT 619
OY 481 CTTTTCGATTTGATGTCTTTATGGAATTCATTAACCTTCACCTTCCTTGATTT 540
DB 620 CTTTTCGATTTGATGTCTTTATGGAATTCATTAACCTTCACCTTCCTTGATTT 679
OY 541 TGTGCTTCAGCTTCGCGCAACAATATCATCTGGGCTGTCTTCTGCGAGGAATGTC 600
DB 680 TGTGCTTCAGCTTCGCGCAACAATATCATCTGGGCTGTCTTCTGCGAGGAATGTC 739
OY 601 ATTGACAAAAGTTAACTGAGGCTGGAAGAACTGAGCTACAAAAGAGAGACAGACT 660
DB 740 ATTGACAAAAGTTAACTGAGGCTGGAAGAACTGAGCTACAAAAGAGAGACAGACT 799

```

```

OY 661 CCACCTATTTAAAGGACCATTTTCAGAAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 720
DB 800 CCACCTATTTAAAGGACCATTTTCAGAAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 859
OY 721 TCCATGTCCTTTAAAGGACCATTTTCAGAAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 780
DB 860 TCCATGTCCTTTAAAGGACCATTTTCAGAAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 919
OY 781 GGATTTCTGTTTGTGCTTTTGTAGTAAATGTTGTAATTTGTCAGATGAGT 840
DB 920 GGATTTCTGTTTGTGCTTTTGTAGTAAATGTTGTAATTTGTCAGATGAGT 979
OY 841 AGTCATGAAGCCTGCTTCATTTTCTCAACATTTCTACTTTTTCATTTACTCTCTT 900
DB 980 AGTCATGAAGCCTGCTTCATTTTCTCAACATTTCTACTTTTTCATTTACTCTCTT 1039
OY 901 TTTTCTTTTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
DB 1040 TTTTCTTTTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1099
OY 961 AAACATGGAATTCG-TTTTGTGTTTACCTTTAGCTCTCTCTCTCTCTCTCTCTCTCT 1019
DB 1100 AAACATGGAATTCG-TTTTGTGTTTACCTTTAGCTCTCTCTCTCTCTCTCTCTCTCT 1159
OY 1020 CACTTATGCTCAATTAATCTTCTAGCAGACATTAATGACATTTACTTTCTATGTTGAA 1079
DB 1160 CACTTATGCTCAATTAATCTTCTAGCAGACATTAATGACATTTACTTTCTATGTTGAA 1219
OY 1080 AAGATTTTTCAAATATGCAATTCGTAATTTGTTAGTTCAGCCTATATTTTGC 1139
DB 1220 AAGATTTTTCAAATATGCAATTCGTAATTTGTTAGTTCAGCCTATATTTTGC 1279
OY 1140 TGGTGGAGATATAGCTGACATTAATCAAGCAATTTTGTGAATTAATGTTT 1199
DB 1280 TGGTGGAGATATAGCTGACATTAATCAAGCAATTTTGTGAATTAATGTTT 1339
OY 1200 CATATGCTTTGTCATTTATAGTTCCTCAAGAACTGCTGGAATTTGTTACTTATTTT 1259
DB 1340 CATATGCTTTGTCATTTATAGTTCCTCAAGAACTGCTGGAATTTGTTACTTATTTT 1399
OY 1260 ACCATTATGCTATTTATAGCTTAAACATTAACCTGCTCCACATCCAGACTGTTTGA 1319
DB 1400 ACCATTATGCTATTTATAGCTTAAACATTAACCTGCTCCACATCCAGACTGTTTGA 1459
OY 1320 ACTGAGTTGCTATGCAATTTGTAATTTCAATTTTACATCTTTTCAACAGACTTT 1379
DB 1460 ACTGAGTTGCTATGCAATTTGTAATTTCAATTTTACATCTTTTCAACAGACTTT 1519
OY 1380 TCAGTGGCCAAATAGTCAGGACATTCAGAAAGTTTATGTTGTAATTCAGTATTTTGA 1439
DB 1520 TCAGTGGCCAAATAGTCAGGACATTCAGAAAGTTTATGTTGTAATTCAGTATTTTGA 1579
OY 1440 ACTGTAAATAGGACTTAAATA-TAGACCATTTTCACAAAGAACAGAAAGAGNGGAA 1498
DB 1580 ACTGTAAATAGGACTTAAATA-TAGACCATTTTCACAAAGAACAGAAAGAGNGGAA 1639
OY 1499 AACATGGAATTTCTTTTGAAGTGCAGCTGTGCTCTCAATTAACATTTTATTTTAAAT 1558
DB 1640 AACATGGAATTTCTTTTGAAGTGCAGCTGTGCTCTCAATTAACATTTTATTTTAAAT 1699
OY 1559 ATATTTTAAATATATGTAAGAAATTAAGTGGCAAGAACTGGAAGACTTAAGACCTGCT 1618
DB 1700 ATATTTTAAATATATGTAAGAAATTAAGTGGCAAGAACTGGAAGACTTAAGACCTGCT 1759
OY 1619 TCAAAAGCCTGGAATATGGAATAAANMNGTTTNCAGATTTTCAT 1667
DB 1760 TCAAAAGCCTGGAATAAATGGAATAAATTTGTTTTCAGATTTTCACAAA 1808

```

RESULT 2
 LOCUS AL578231/c 974 bp mRNA linear EST 01-JUN-2003
 DEFINITION AL578231 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens

ACCESSION cDNA clone CS0DK002Y120 3-PRIME, mRNA sequence.
 AL578231
 VERSION AL578231.2 GI:31316449
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 974)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 16, 2001 this sequence version replaced gi:12942110.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6664.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DK002BE10NP1;cluster=6664.f. Contact :
 Peng Liang Email : filiang@life.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DK002BE10NP1.
 Location/Qualifiers

FEATURES
 source
 1..974
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK002Y120"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 344 a 163 c 170 g 263 t 34 others
 ORIGIN

Query Match 47.08; Score 872.4; DB 9; Length 974;
 Best Local Similarity 91.88; Pred. No. 4,7e-127;
 Matches 893; Conservative 29; Mismatches 50; Indels 1; Gaps 1;
 682 GCAGATTGCAAAAAATCTTCAGTCTTTGGCTATTCAGTCTTAAACTTG 741
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 974 GCAGATTGCAAAAAATCTTCAGTCTTTGGCTATTCAGTCTTAAACTTG 915
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 742 AGTAGCTTTCTGTTGACTTGGCCCTACATCCTTCGGATTCTGTTTGTCTT 801
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 914 AGTAGCTTTGCTTCTGACTTGGCCCTACACCTTCGGATTCTGTTTGTCTT 855
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 802 GTAGTAGTAAAGTGTGAATGTTATGGCGATCGAGTAGTCATGAAGCCTGTCAT 861
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 854 GTAGTAGTAAAGTGTGAATGTTATGGCGATCGAGTAGTCATGAAGCCTGTCAT 795
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 862 TTTCCTCAACTATCTCTCTTTTCACTTACTCTCTTTTCTTTTCTTCCACCTCG 921
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 794 TTTCCTCAACTATCTCTCTTTTCACTTACTCTCTTTTCTTTTCTTCCACCTCG 735
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 922 TTTCCTCAACTATCTCTCTTTTCACTTACTCTCTTTTCTTTTCTTCCACCTCG 981
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 734 TTTCCTCAACTATCTCTCTTTTCACTTACTCTCTTTTCTTTTCTTCCACCTCG 675
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 982 GTGTTACCTTACTCTCTCTTTTCACTTACTCTCTTTTCTTTTCTTCCACCTCG 1041
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 674 GTGTTACCTTACTCTCTCTTTTCACTTACTCTCTTTTCTTTTCTTCCACCTCG 615
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 1042 CTAGCAGACAAATAGACATTAATCTTATGTTGGAAAAAGATTTCACAAAGTATGCA 1101
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 614 CTAGCAGACAAATAGACATTAATCTTATGTTGGAAAAAGATTTCACAAAGTATGCA 555

QY 1102 ATTCTGAATAATTTAGTTCAGCTCATATATTTCTGTTGGATATAGCTACTCA 1161
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 554 ACAGAAAAATATTTAGTTCAGCTCATATATTTCTGTTGGATATAGCTACTCA 495
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1162 TTGAATTCACCAATTTTGGAAATTTAATGTTTTCATATCTGTTTCATTGTA 1221
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 494 TTGAATTCACCAATTTTGGAAATTTAATGTTTTCATATCTGTTTCATTGTA 435
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1222 GTTCCAGAAAGCTGGAATTTGTTTCTTCAATTTTACCTTATGATATATAGGCTT 1281
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 434 GTTCCAGAAAGCTGGAATTTGTTTCTTCAATTTTACCTTATGATATATAGGCTT 375
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1282 AACATTAATCTGCTCCACATCCAGACTGTTTGTGAATGATGCTATGCAATGTT 1341
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 374 AACATTAATCTGCTCCACATCCAGACTGTTTGTGAATGATGCTATGCAATGTT 315
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1342 AATTCAATATTTTACATCTTTCTGAAACAGACTTTTCAGTGGCCAAATAGTCAGAC 1401
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 314 AATTCAATATTTTACATCTTTCTGAAACAGACTTTTCAGTGGCCAAATAGTCAGAC 255
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1402 ATTCAAGGTTTATGTTGTAATATCATGATATTTTGAACCTGTAATAATGACTTAATA 1461
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 254 ATTCAAGGTTTATGTTGTAATATCATGATATTTTGAACCTGTAATAATGACTTAATA 195
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1462 -TAGACCATTTCTACAAAGAACCACTGAATAGGAAACATGGAATTTCTTTAGGTG 1520
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 194 TTACACCATTTCTACAAAGAACCACTGAATAGGAAACATGGAATTTCTTTAGGTG 135
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1521 CAGTGTGCTCTTAAATTTACATTTAGTTTATATATTTTAAACATATGTAAGAA 1580
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 134 CAGTGTGCTCTTAAATTTACATTTAGTTTATATATTTTAAACATATGTAAGAA 75
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1581 ATTAAGTGGCAAGACCTGGAAGCTTAAGACCTGCTCAAGACCTGAATATAGGAA 1640
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 74 ATTAAGTGGCAAGACCTGGAAGCTTAAGACCTGCTCAAGACCTGAATATAGGAA 15
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1641 AATAAANWNGTGT 1653
 ||:||||:||||
 DB 14 AVDAAGKGGTNT 2

RESULT 3
 AL555819 958 bp mRNA linear EST 31-MAY-2003
 LOCUS AL555819
 DEFINITION Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 ACCESSION cDNA clone CS0DK002Y120 5-PRIME, mRNA sequence.
 VERSION AL555819
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 958)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12897913.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6664.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DK002BE10NP1;cluster=6664.f. Contact :
 Peng Liang Email : filiang@life.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DK002BE10NP1.
 Location/Qualifiers
 1..958
 /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK002Y120"
/cell_line="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/Note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      221 a      213 c      206 g      310 t      8 others
ORIGIN

```

```

Query Match      40.9%; Score 760; DB 9; Length 958;
Best Local Similarity 97.5%; Pred. No. 1.9e-109;
Matches 780; Conservative 2; Mismatches 17; Indels 1; Gaps 1;

QY 10 CTAGAGGCTTACTGTTCTCGGCGCCCTTGAGCTTACTTTTGTCTGCTGCTCTC 69
DB 160 STGGAAGGTACTATTCTCGCGCCCTTGAGCTTACTTTTGTATCTGCTCTC 219
QY 70 TTCTCCGCTTACGCGCGCGCTGAGAGCCCTACATGAGAGATCTTCCACCTGCT 129
DB 220 TTCTCCGCTTACGCGCGCGCTGAGAGCCCTACATGAGAGATCTTCCACCTGCT 279
QY 130 CAGGCGAGCGCTACTGTAGAGGCCATTCTCCCTTCCAGTGGATCCCATGATTA 189
DB 280 CAGGCGCA-SGCTACTGTAGAGGCCATTCTCCCTTCCAGTGGATCCCATGATTA 338
QY 190 ACATTACTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
DB 339 ACATTACTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 398
QY 250 GGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
DB 399 GGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
QY 310 AGTGTGCAACCTTCTATTACTATTATTGCTTTTCCACAGATCAACCCAGAAACAG 369
DB 459 AGTGTGCAACCTTCTATTACTATTATTGCTTTTCCACAGATCAACCCAGAAACAG 518
QY 370 GCTGCTCAAGATTCAGAGAGCTGTCACATTAACATGAGATTTTCCACACTT 429
DB 519 GCTGCTCAAGATTCAGAGAGCTGTCACATTAACATGAGATTTTCCACACTT 578
QY 430 TATTTTAACTTCTTATTATACAGAGAGATCTATGTTTAACTTCTTATGCA 489
DB 579 TATTTTAACTTCTTATTATACAGAGAGATCTATGTTTAACTTCTTATGCA 638
QY 480 TATTTGATGTCCTTTATGAAATCATAAACTCAGCCTTCTTGATTTTGTGCTTC 549
DB 639 TATTTGATGTCCTTTATGAAATCATAAACTCAGCCTTCTTGATTTTGTGCTTC 698
QY 550 ATGTTGGGCAAAACAATATATCTGGGCTGCTCTGAGGAGAAATGCTATTCACAA 609
DB 699 ATGTTGGGCAAAACAATATATCTGGGCTGCTCTGAGGAGAAATGCTATTCACAA 758
QY 610 AAGTTAACTGAGGCTTGGAAAACTGAGCTACAAAGAAGAGAGACTTCCACCTAT 669
DB 759 AAGTTAACTGAGGCTTGGAAAACTGAGCTACAAAGAAGAGAGACTTCCACCTAT 818
QY 670 AAAGGACATTTGAGAAATCAGAAAAATTTCTCAGTTTCTTTGGCTTATTCATG 729
DB 819 AAAGGACATTTGAGAAATCAGAAAAATTTCTCAGTTTCTTTGGCTTATTCATG 878
QY 730 TTTAAAACTGAGATAGCTTTTCTGTTGACTTGGCCCTACATCTTCTGGGATTTCTG 789
DB 879 TTTAAAACTGAGATAGCTTTTCTGTTGACTTGGCCCTACATCTTCTGGGATTTCTG 938
QY 790 TTTTGTGCTTTGTAGTAGT 809
DB 939 TTTTGTGCTTTGTAGTAGT 958

```

```

RESULT 4
LOCUS      B1770281
DEFINITION 603055973P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5205402 5',
            mRNA sequence.
ACCESSION  B1770281
VERSION    B1770281.1 GI:15761859
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 784)
            NIH-MGC http://mgi.cni.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1515 row: a column: 19
            High quality sequence stop: 756.
FEATURES
    source
        1..784
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5205402"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_122"
        /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
        Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
        anonymous pool of 24 week female lung, 16 week female
        spleen, and 20-22 week male spleens. Library is oligo-dT
        primed and directionally cloned (EcoRV site is destroyed
        upon cloning). Average insert size 1.4 kb, insert size
        range 1-3 kb. Library is normalized and enriched for
        full-length clones and was constructed by C. Gruber
        (Invitrogen). Research Genetics tracking code 026. Note:
        this is a NIH-MGC library."
BASE COUNT      180 a      187 c      170 g      247 t
ORIGIN
Query Match      34.4%; Score 639; DB 12; Length 784;
Best Local Similarity 97.6%; Pred. No. 1.8e-90;
Matches 659; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1 ATGGGCGAGCTAGAGGTTACTGTTTCGCGCGGCTGAGCTTATTTAGTGC 60
DB 87 ATGGGCGAGCTAGAGGTTACTGTTTCGCGCGGCTGAGCTTATTTAGTGC 146
QY 61 TGCCCTCTCTCTCCGCTTACAGCGGCGCTGAGAGCCCTACATGAGAGATCTTC 120
DB 147 TGCCCTCTCTCTCCGCTTACAGCGGCGCTGAGAGCCCTACATGAGAGATCTTC 206
QY 121 CACCTGCTCAGGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATGCC 180
DB 207 CACCTGCTCAGGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATGCC 266
QY 181 ATGATTACTACTATTACCTGCGCTTGTACTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 267 ATGATTACTACTATTACCTGCGCTTGTACTGCTGCTGCTGCTGCTGCTGCTGCT 326
QY 241 TGATCTTTGATGCTGAGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 327 TGATCTTTGATGCTGAGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 386

```


QY 301 CTTCTCTCAGTGTGGCAACTCTATTAATTAATTTGCTTTCCAAAGGTACAAACC 360
 Db 387 CTTCTCTCAGTGTGGCAACTCTATTAATTAATTTGCTTTCCAAAGGTACAAACC 446
 QY 361 AGAAACAAGGCTCCCTCAATATCCAGAGAGTCTTGCAACATTACACATGACATTTT 420
 Db 447 AGAAACAAGGCTCCCTCAATATCCAGAGAGTCTTGCAACATTACACATGACATTTT 506
 QY 421 CCAACACTTATTTTATTTTAACTCTCTTATTAATACAGAGAGTATGTTTTTACT 480
 Db 507 CCAACACTTATTTTCTTTAACTCTCTTATTAATACAGAGAGTATGTTTTTACT 566
 QY 481 CTTTTCATATTTTATGATGTCTTTATGAAATCATATAAACTTCCAGCTTCTTGATTT 540
 Db 567 CTTTTCATATTTTATGATGTCTTTATGAAATCATATAAACTTCCAGCTTCTTGATTT 626
 QY 541 TGTGGCTCATGTTTGGGCAAAATATATCTGGGCTCTTCTTGTGAGGA-ATGT 599
 Db 627 CGTGGCTCATGTTTGGGCAAAATATATCTGGGCTCTTCTTGTGAGGAATGT 686
 QY 600 CATTCGCAAAAGTTAACTGAGCTTGAAAACTGAGCTCAAAAGAGAGACAGACT 659
 Db 687 CATTCGCAAAAGTTAACTGAGCTTGAAAACTGAGCTCAAAAGAGAGACAGACT 746
 QY 660 TCCACCTATTAAAG 674
 Db 747 TCCACCTATTAAAG 761

RESULT 5
 B1825977 821 bp mRNA linear EST 04-OCT-2001
 LOCUS 60307650F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5168319 5',
 DEFINITION mRNA sequence.

ACCESSION B1825977
 VERSION B1825977.1 GI:15937527
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 821)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.lnl.gov
 Plate: L14M1418 row: h column: 16
 High quality sequence start: 6
 High quality sequence stop: 800.

FEATURES

location/Qualifiers
 1..821
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5168319"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_119"
 /note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber

(Invertegen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC library.
 BASE COUNT 187 a 195 c 179 g 260 t

Query Match 34.2%; Score 634.2; DB 12; Length 821;
 Best Local Similarity 97.4%; Pred. No. 1e-89;
 Matches 708; Conservatve 0; Mismatches 13; Indels 6; Gaps 6;

QY 1 ATGGCGAGCTAAGGTTACTGTCTCGGGCCCTTGAGCGTACCTTTTAGTGC 60
 Db 101 ATGGCGAGCTAAGGTTACTGTCTCGGGCCCTTGAGCGTACCTTTTAGTGC 160
 QY 61 TGCCTCCTTTCGCCGCTTCAGCGGGCCCTGAGAGCCCTACATGAGAGATCTC 120
 Db 161 TGCCTCCTTTCGCCGCTTCAGCGGGCCCTGAGAGCCCTACATGAGAGATCTC 219
 QY 121 CACCTGCTCAGGCGCAGCCCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATCC 180
 Db 220 CACCTGCTCAGGCGCAGCCCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATCC 279
 QY 181 ATGATTACTATTAATCTGCTGTACTGTGTCACTGTGAGTGGTCAAACTGCCATT 240
 Db 280 ATGATTACTATTAATCTGCTGTACTGTGTCACTGTGAGTGGTCAAACTGCCATT 339
 QY 241 TGGATCTTTGATGTCGTAACAGTGTGCTGCCATTTGGGATGCTAGATTGTTAAT 300
 Db 340 TGGATCTTTGATGTCGTAACAGTGTGCTGCCATTTGGGATGCTAGATTGTTAAT 397
 QY 301 CTTCTCTCAGTGTGGCAACTCTATTAATTAATTTGCTTTCCAAAGGTACAAACC 360
 Db 398 CTTCTCTCAGTGTGGCAACTCTATTAATTAATTTGCTTTCCAAAGGTACAAACC 457
 QY 361 AGAAACAAGGCTCCCTCAATATCCAGAGAGTCTTGCAACATTACACATGACATTTT 420
 Db 458 AGAAACAAGGCTCCCTCAATATCCAGAGAGTCTTGCAACATTACACATGACATTTT 517
 QY 421 CCAACACTTATTTTATTTTAACTCTCTTATTAATACAGAGAGTATGTTTTTACT 480
 Db 518 CCAACACTTATTTTCTTTAACTCTCTTATTAATACAGAGAGTATGTTTTTACT 577
 QY 481 CTTTTCATATTTTATGATGTCTTTATGAAATCATATAAACTTCCAGCTTCTTGATTT 540
 Db 578 CTTTTCATATTTTATGATGTCTTTATGAAATCATATAAACTTCCAGCTTCTTGATTT 635
 QY 541 TGTGGCTCATGTTTGGGCAAAATATATCTGGGCTCTTCTTGTGAGGAATGT 600
 Db 636 TGTGGCTCATGTTTGGGCAAAATATATCTGGGCTCTTCTTGTGAGGAATGT 695
 QY 601 ATTGCACAAAGTTAACTGAGCTTGAAAACTGAGCTACAAAGAGAGACAGACT 660
 Db 696 ATTGCACAAAGTTAACTGAGCTTGAAAACTGAGCTACAAAGAGAGACAGACT 755
 QY 661 CCACTATTAAAGGACCATTTGAGATTCAGAAAAATTTCTTCAAGTTCTTTGGCTTAT 720
 Db 756 CCACTATTAAAGGACCA-TTGGCGAATTCAGAAAAATTTCTTCAAGTTCTTTGGCTTAT 814
 QY 721 TCCATGT 727
 Db 815 TCCATGT 821

RESULT 6
 CD252003 911 bp mRNA linear EST 22-MAY-2003
 LOCUS AGENCOURT_14211757 NIH_MGC_180 Homo sapiens cDNA clone
 DEFINITION IMAGE:30383501 5', mRNA sequence.
 ACCESSION CD252003
 VERSION CD252003.1 GI:31012469
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 911)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: NDAM47 row: h column: 06
 High quality sequence stop: 665.
 Location/Qualifiers
 1..911

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone IMAGE:30383501"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_lib="NIH_MGC_180"
 /note="Organ: Testis; Vector: PCMV-SPORT.1; Site.1: NotI; Site.2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH-MGC Library."
 location/Qualifiers
 1..911

BASE COUNT
 196 a 211 c 199 g 305 t

Query Match 33.3%; Score 618.6; DB 14; Length 911;
 Best Local Similarity 90.2%; Pred. NO. 2.7e-87;
 Matches 708; Conservative 0; Mismatches 44; Indels 33; Gaps 3;

1 ATGGCGCAGTAGAGGGTTACTGTTTCGGCCGCTTGAGCTGACCTTTTAACTGTC 60
 128 ATGGCGCAGTAGAGGGTTACTGTTTCGGCCGCTTGAGCTGACCTTTTAACTGTC 187
 61 TGCCCTCTCTTCGCGCTTCAGCGCGGCGCTGCGAGAGCCCTACAGAGAGATCTTC 120
 188 TGCCCTCTCTTCGCGCTTCAGCGCGGCGCTGCGAGAGCCCTACAGAGAGATCTTC 247
 121 CACCTGCTCAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 248 CACCTGCTCAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281
 181 ATGATTACTATTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 282 ATGATTACTATTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 341
 241 TGGATCTTGGATGGTGTGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 342 TGGATCTTGGATGGTGTGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
 301 CTTCCTCTCACTGTTGGCACTTCTGATTACTATTATTTGCTTTCCAAAGGTACACC 360
 402 CTTCCTCTCACTGTTGGCACTTCTGATTACTATTATTTGCTTTCCAAAGGTACACC 461
 361 AGAACAAGGCTGCTCAAGTATCAGAGAGCTTGTACATTAACACTAGCACTATTT 420
 462 AGAACAAGGCTGCTCAAGTATCAGAGAGCTTGTGTAACTTACACTAGCACTATTT 521
 421 CCAACACTTATTTTAACTCTCTTATTTATACAGAGAGATCTATGTTTAACT 480
 522 CCAACACTTATTTTAACTCTCTTATTTATACAGAGAGATCTATGTTTAACT 581
 481 CTTTTCATATTTGATGTGCTTTATGGAATCATTAACCTCAGCCTTCCTGGATTT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 582 CTTTTCATATTTGATGTGCTTTATGGAATCATTAACCTCAGCCTTCCTGGATTT 641
 541 TGTGGCTTCATGTTTCGGCAACAATATCATTCATTCGGCTGCTCTGTGACGGAATGTC 600
 642 TGTGGCTTCATGTTTCGGCAACAATATCATTCATTCGGCTGCTCTGTGACGGAATGTC 701
 601 ATTGCACAAAGTTTAACTGAGCTTGGCAAACTGACCTACAAAGAAGAGACACTT 660
 702 ATTGCACAAAGTTTAACTGAGCTTGGCAAACTGACCTACAAAGAAGAGACACTT 761
 661 --CCACCTATTAAGGACCAATTTGCAGATTCAGAAAAA-----TTCCTGAGTTCTTTT 713
 762 TCCACTATTTAAAGGACCAATTTGCAGATTCAGAAAAAATCTTTCAGTTTCTTTTGG 821
 714 GGCCTTATTCATGCTCTTTTAAAGCACTTGAGTATGCTTTTCTGTTGACTTGGCCCTACAT 773
 822 GCTTATTTCCATGCTCTTTTAAAGCACTTGAGTATGCTTTTCTGTTGACTTGGCCCTACAT 881
 774 CCTTC 778
 |||||
 882 CCTAC 886

RESULT 7
 BE738701/c 1019 bp mRNA linear EST 15-SEP-2000
 LOCUS 60157294471 NIH_MGC_57 Homo sapiens CDNA clone IMAGE:3839674 3,
 DEFINITION mRNA sequence.
 BE738701
 BE738701.1 GI:10152693
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1019)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>.
 Plate: LDCM528 row: h column: 11
 High quality sequence start: 7
 High quality sequence stop: 721.
 Location/Qualifiers
 1..1019

FEATURES
 source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone IMAGE:3839674"
 /clone_lib="NIH_MGC_57"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site.1: SfiI (99ccgctcggcc); Site.2: SfiI (99ccatctggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTTAGAGCCGAGGCGCCGACATG-(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
 BASE COUNT
 395 a 168 c 177 g 279 t

Query Match	32.1%;	Score 596;	DB 10;	Length 1019;
Best Local Similarity	89.5%;	Pred. No. 9.4e-84;		
Matches 719;	Conservative 0;	Mismatches 76;	Indels 8;	Gaps 7;
QY	798	TTTTGAGTAGTTAATAGTGAATTTTATGCGCATGCGAGTAGTCATGAACCTGTCT	857	
DB	802	TGCTCTTGTATATATATGTTGGTATGTCATCTAGTACTCATGAAGTCCGTC	743	
QY	858	TCATTTTCTCAACATTTCTACTTTTTCATTTTCTCTCTTTTCTTCTCTCTCT	917	
DB	742	CTAATGTTCTTCACTAATTAACGTTTTCATTTCTCTCTCTCTCTCTCTCTCT	684	
QY	918	CCTGTCCTCAGCAAAATTTAGACTTTTCTCTCTCTCTCTCTCTCTCTCTCT	976	
DB	683	CCGTCTCCTCAGCAAAATTTAGACTTTTCTCTCTCTCTCTCTCTCTCTCTCT	624	
QY	977	TTTTGGTGTACCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1034	
DB	623	TTTTTGTGTACCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	564	
QY	1035	ATACCTTGTAGAGAGCAATAGACA--TTATCTTCTAGTGTGGAAGAAGTTTTC	1093	
DB	563	ATACCTTGTAGAGAGCAATAGACAATTTATCTTCTCTCTCTCTCTCTCTCT	504	
QY	1094	GATATGCAATTCGAAATTTTGTAGTTCCAGCCTATATATTGCTGTGAGATAG	1153	
DB	503	GATATGCAATTCGAAATTTTGTAGTTCCAGCCTATATATTGCTGTGAGATAG	444	
QY	1154	CTGACTCATTTGAATCAAGCCAAATTTTGAATTAATGTTTTCATATCTCTGT	1213	
DB	443	CTGACTCATTTGAATCAAGCCAAATTTTGAATTAATGTTTTCATATCTCTGT	384	
QY	1214	TTGTATAGTTCCTCAGAAAGCTGGAATTTCTTACTTCTATCTTATCTTAT	1273	
DB	383	CTGTATAGTTCCTCAGAAAGCTGGAATTTCTTACTTCTATCTTATCTTAT	324	
QY	1274	ATAGGCTTAAACATACTCTGCTCCACATCCAGACTTGTGTGAACTGAGTGTAT	1333	
DB	323	ATAGGCTTAAACATACTCTGCTCCACATCCAGACTTGTGTGAACTGAGTGTAT	264	
QY	1334	CAATGTTTAAATTTCACTTTTTCATCTTCTCTGACAGACT--TTTCAGTGGCC	1392	
DB	263	CAGTGTAAATTTCACTTTTTCATCTTCTCTGACAGACTTTCAGTGGCCAA	204	
QY	1393	AGTCAGGACATCAAGGTTTATGTTGTAATTCAGTGTATTTGAACGTAAAA	1452	
DB	203	AGTCAGGACATCAAGGTTTATGTTGTAATTCAGTGTATTTGAACGTAAAA	144	
QY	1453	ACTTAATTA--TAGACATTTTCAACAAGCAACTGAATAGGNGGAAACATGA	1511	
DB	143	ACTTAATTAATTTAGACATTTTCAACAAGCAAGCAATAGTGGAACATGA	84	
QY	1512	TTTTAGGTGACAGTGTGCTTCAAAATTAACATTTTCTT--TTAATATATAT	1570	
DB	83	TTTTAGGTGACAGTGTGCTTCAAAATTAACATTTTCTTCTATATATAT	24	
QY	1571	TATGTAGAAATTAAGTGCAAA 1593		
DB	23	TATGTAGAAATTAAGTGCTACA 1		
RESULT 8				
AL519606	1201 bp	mRNA	linear	EST 22-MAY-2003
DEFINITION	AL519606	Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED	Homo sapiens	
ACCESSION	AL519606	CDNA clone CS0DB004Y06 5-PRIME, mRNA sequence.		
VERSION	AL519606.2	GI:31037953		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.				
AUTHORS	1 (bases 1 to 1201)				
TITLE	Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished On Feb 13, 2001 this sequence version replaced gi:12783099. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6664.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DB004DE03QPI&cluster=6664.f. Contact : Feng Liang Email: fliang@life.techn.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DB004DE03QPI. Location/Qualifiers				
FEATURES	1..1201				
SOURCE	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DB004Y06" /issue_type="NEUROBLASTOMA COT 10-NORMALIZED" /clone_1id="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."				
BASE COUNT	370 a 185 c 223 g 360 t 63 others				
ORIGIN					
Query Match	31.3%; Score 580.4; DB 9; Length 1201;				
Best Local Similarity	92.1%; Pred. No. 2.5e-81;				
Matches 598;	Conservative 8; Mismatches 42; Indels 1; Gaps 1;				
QY	1106	TGAATATTTTGTAGTTCAGGCTATATTTTGTGCTGTTGAGATAGCTGATCTGA	1165		
DB	66	TAAATATTTTGTAGTTCAGGCTATATTTTGTGCTGTTGAGATAGCTGATCTGA	125		
QY	1166	AATCAAGCCAAATTTTGTGAATTAATGTTTTCATATGCTGTTCAATGTTAATG	1225		
DB	126	AATCAAGCCAAATTTTGTGAATTAATGTTTTCATATGCTGTTCAATGTTAATG	185		
QY	1226	CTCAGAAAGCTGCTGGAATTTCTTACTTATTTTACCTATATGTCATTTTACG	1285		
DB	186	CTCAGAAAGCTGCTGGAATTTCTTACTTATTTTACCTATATGTCATTTTACG	245		
QY	1286	TAACTCTGCTCCACATCCAGACTGTTTGTGAAGTGTGCTATGCAATTTGTA	1345		
DB	246	TAACTCTGCTCCACATCCAGACTGTTTGTGAAGTGTGCTATGCAATTTGTA	305		
QY	1346	TCATTAATTTTTCATCTTTCGACAGACTTTTCAGTGGCCAAATTAAGTGA	1405		
DB	306	TCATTAATTTTTCATCTTTCGACAGACTTTTCAGTGGCCAAATTAAGTGA	365		
QY	1406	AAAGGTTTATGTTGTAATTCAGTGTATTTTGAAGTGAAGTGAAGTGAAGT	1464		
DB	366	AAAGGTTTATGTTGTAATTCAGTGTATTTTGAAGTGAAGTGAAGTGAAGT	425		
QY	1465	ACCATTTTCAACAAGCAACTGAATAGGNGGAAACATGAATTTCTTTAGT	1524		
DB	426	ACCATTTTCAACAAGCAACTGAATAGTGAAGTGAAGTGAAGTGAAGTGA	485		
QY	1525	GCTGCTCTTCAAAATTAACATTTTCTTCTATATATTTTAAACATATGA	1584		
DB	486	GCTGCTCTTCAAAATTAACATTTTCTTCTATATATTTTAAACATATGA	545		
QY	1585	ACTGCAAGAACTGGAAGCTTAAGACCTCTTCAANAGCTGATTAATGGA	1644		
DB	546	ACTGCAAGAACTGGAAGCTTAAGACCTCTTCAANAGCTGATTAATGGA	605		

BASE COUNT 131 a 169 c 148 g 219 t
ORIGIN

Query Match 27.4%; Score 509.4; DB 10; Length 667;
Best Local Similarity 97.9%; Pred. No. 3.7e-70;
Matches 516; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGGCGACAGTGAAGGATTACTGTTCTGCGCGCCCTTGAGCTGATCTTTTATGATGCC 60
|||||
Db 141 ATGGCGACAGTGAAGGATTACTGTTCTGCGCGCCCTTGAGCTGATCTTTTATGATGCC 200
QY 61 TGGCTCTCTCTTCCGCTTCAAGCCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 120
|||||
Db 201 TGGCTCTCTCTTCCGCTTCAAGCCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 260
QY 122 CACCTGCTCAGGCGAGGCTGATGAGGGGCAATTCCTCCCTCCCAAGTGGATCCC 180
|||||
Db 261 CACCTGCTCAGGCGAGGCTGATGAGGGGCAATTCCTCCCTCCCAAGTGGATCCC 320
QY 181 ATGATTACTACATTACCTGCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
|||||
Db 321 ATGATTACTACATTACCTGCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
QY 241 TGGATCTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
|||||
Db 381 TGGATCTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440
QY 301 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
|||||
Db 441 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500
QY 361 AGAACAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
|||||
Db 501 AGAACAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
QY 421 CCAACACTTATTTTAACTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 480
|||||
Db 561 CCAACACTTATTTTAACTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 620
QY 481 CTTTTCGATTTTGTGCTGCTTATTTGAAATCATTAACCTCAGC 527
|||||
Db 621 CTTTTCGATTTTGTGCTGCTTATTTGAAATCATTAACCTCAGC 667

RESULT 11
B1768769 694 bp mRNA linear EST 25-SEP-2001
LOCUS 60305826F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5207486 5',
DEFINITION mRNA sequence.
ACCESSION B1768769
VERSION B1768769.1 GI:15760347
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1520 row: h column: 15
High quality sequence stop: 694.
Location/Qualifiers
1. 694

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5207486"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: PCMV-SPOrt6;
Site.1: NotI; Site.2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 154 a 155 c 164 g 220 t 1 others
ORIGIN

Query Match 27.1%; Score 503.4; DB 12; Length 694;
Best Local Similarity 94.0%; Pred. No. 3.2e-69;
Matches 611; Conservative 0; Mismatches 27; Indels 12; Gaps 8;

QY 1 ATGGCGACAGTGAAGGATTACTGTTCTGCGCGCCCTTGAGCTGATCTTTTATGATGCC 60
|||||
Db 47 ATGGCGACAGTGAAGGATTACTGTTCTGCGCGCCCTTGAGCTGATCTTTTATGATGCC 105
QY 61 TGGCTCTCTCTTCCGCTTCAAGCCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 119
|||||
Db 106 TGGCTCTCTCTTCCGCTTCAAGCCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 165
QY 120 CCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 178
|||||
Db 166 CCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225
QY 179 CCATGATTTACTACATTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238
|||||
Db 226 CCATGATTTACTACATTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
QY 239 TTGGGATTTGGATGCTGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 298
|||||
Db 286 TTGGGATTTGGATGCTGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
QY 299 ATCTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
|||||
Db 345 ATCTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
QY 359 CCAAGAAACAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
|||||
Db 405 CCAAGAAACAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464
QY 418 TTTCGCAACCTTATTTT -TAACCTCTTATTTATTCAGAGAGAGATCTTATTTT 475
|||||
Db 465 TTTCGCAACCTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 524
QY 476 TTACTC-----TTTTCATATTTGATGCTGCTTATGAGAAATATTAACCTGAGCTTC 531
|||||
Db 525 TGCCTCTGATATTTGCTATTTGATGCTGCTTATGAGAAATATTAACCTGAGCTTC 584
QY 532 CTTGGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
|||||
Db 585 CTTGGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
QY 592 GGAATGCTATTCACAAAGTAACTGAGGCTTGGAAAACGAGACTAC 640
|||||
Db 645 GGAATGCTATTCACAAAGTAACTGAGGCTTGGAAAACGAGACTAC 694

RESULT 12
A0599053/c 560 bp DNA linear GSS 08-JUN-1999
LOCUS HS_2133_B2.B11.T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=2133 Col=22 Row=D, genomic survey

sequence.
 ACCESSION AO39053.1 GI:5021428
 VERSION AO39053.1
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 560)
 MAHAIRAS G.G., WALLACE J.C., SMITH R., SWARTZELL S., HOLZMAN T., KELLER A., SHAKER R., FURLONG J., YOUNG J., ZHAO S., ADAMS M.D. and HOOD L.,
 Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 JOURNAL MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallaceu.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.husc.washington.edu
 Plate: 2133 Row: D Column: 22
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 560.

FEATURES
 Source Location/Qualifiers
 1..560
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="plate=2133 Col=22 Row=D"
 /sex="male"
 /clone_id="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: Sperm; Vector: pBelBAC11; BAC clones in E-Coli DH10B"

BASE COUNT 195 a 103 c 85 g 171 t 6 others

ORIGIN
 Query Match 25.3%; Score 469.8; DB 28; Length 560;
 Best Local Similarity 98.0%; Pred. No. 6.1e-64;
 Matches 493; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 1107 GAATATTTGTTAGTCCAGCTATATATTCGTGCTGGATAGCTGACATTTGAA 1166
 |||||
 DB 559 GAAATATTTGTTAGTCCAGCTATATATTCGTGCTGGATAGCTGACATTTGAA 500
 |||||
 QY 1167 ATCAAGCAATTTTGGATTTAATGTTTTCATATGCTGTTCAATTTGTAATAGTCC 1226
 |||||
 DB 499 ATCAAGCAATTTTGGATTTAATGTTTTCATATGCTGTTCAATTTGTAATAGTCC 440
 |||||
 QY 1227 TCAGAAAGCTGTGAATTTGCTTACTTCATTTACCTTATGTCATTTATAGCGTTAAACAT 1286
 |||||
 DB 439 TCAGAAAGCTGTGAATTTGCTTACTTCATTTACCTTATGTCATTTATAGCGTTAAACAT 380
 |||||
 QY 1287 AACTCTGCTCCACATCCAGACTGTTTGGAACTGATGTCATGCAATTTGTAATTT 1346
 |||||
 DB 379 AACTCTGCTCCACATCCAGACTGTTTGGAACTGATGTCATGCAATTTGTAATTT 320
 |||||
 QY 1347 CATACACTTTTACATCTTTTGAACAAGACTTTCAAGTGGCCAAATGTCAGACATTTCA 1406
 |||||
 DB 319 CATACACTTTTACATCTTTTGAACAAGACTTTCAAGTGGCCAAATGTCAGACATTTCA 260
 |||||
 QY 1407 AAGCTTATGTGTAATATCATGATATTTTGAAGCTGTAAAAATGACTTAATAA-TAGA 1465
 |||||
 DB 259 AAGCTTATGTGTAATATCATGATATTTTGAAGCTGTAAAAATGACTTAATAA-TAGA 200
 |||||
 QY 1466 CCATTTCACAAAGAACATGATAGTGGAAACATGGAATTTCTTTTAGGTGCACTG 1525
 |||||

DB 199 CCATTTCACAAAGAACATGATAGTGGAAACATGGAATTTCTTTTAGGTGCACTG 140
 QY 1526 GTGGCTTCACAAATTTACATTTGTTTAAATATATATTTTAAACATATGTAAGAAATTTAA 1585
 |||||
 DB 139 GTGGCTTCACAAATTTACATTTGTTTAAATATATATTTTAAACATATG-TAG-AAGAAATTTAA 81
 |||||
 QY 1586 GTGGCAAGAACTGGGAAGCTT 1608
 |||||
 DB 80 GTGGCAAGAACTGGGAAGCTT 58
 |||||

RESULT 13
 BX401726/c 1126 bp mRNA linear EST 13-MAY-2003
 LOCUS BX401726 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DL007YH07 3-PRIME, mRNA sequence.
 ACCESSION BX401726
 VERSION BX401726.1 GI:30618581
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1126)
 LI W.B., GRUBER C., JESSEE J. and POLYES D.
 Full-length cDNA libraries and normalization
 Unpublished
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6666.f,
 Contact : Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DL007C04NP1.

FEATURES
 source Location/Qualifiers
 1..1126
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL007YH07"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_id="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 390 a 202 c 210 g 314 t 10 others

ORIGIN
 Query Match 25.0%; Score 464; DB 13; Length 1126;
 Best Local Similarity 96.9%; Pred. No. 4.4e-63;
 Matches 473; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 160 TCCCTTCCAGTGGATGCCATGATTTACATTTACCTGCTGTACTGCTGCTAGTT 219
 |||||
 DB 497 TTTATTTTAAAGTGGATGCCATGATTTACATTTACCTGCTGTACTGCTGCTAGTT 438
 |||||
 QY 220 GGAATGTCAAACCTGCCATTTGGATCTTTGGATGTCGTAACATGTTGTGCTGCATTT 279
 |||||
 DB 437 GGAATGTCAAACCTGCCATTTGGATCTTTGGATGTCGTAACATGTTGTGCTGCATTT 378
 |||||
 QY 280 GGGATGTCGATTTGTTAATCTTCTCTGCTGCTGGCACTCTATTTACTATATTTT 339
 |||||
 DB 377 GGGATGTCGATTTGTTAATCTTCTCTGCTGCTGGCACTCTATTTACTATATTTT 318
 |||||
 QY 340 CTTTTCACAAAGTATACCAAGCAAGAGCTCCTCAAGTATCCAGAGAGTCTTGCA 399
 |||||
 DB 317 CTTTTCACAAAGTATACCAAGCAAGAGCTCCTCAAGTATCCAGAGAGTCTTGCA 258
 |||||

QY 400 ACATTAACTACAGATATTTCCACACTTTATTTTAACTTCCTTTATATATACGAA 459
 Db 257 ACATTAACTACAGATATTTCCACACTTTATTTTAACTTCCTTTATATATACGAA 138
 QY 460 GCAGGATCTATGTTTTTACTCTTTTGCATATTTGATGCTCTTATGGAATCATATA 519
 Db 197 GCAGGATCTATGTTTTTACTCTTTTGCATATTTGATGCTCTTATGGAATCATATA 138
 QY 520 ACATTCACCTTCCTGATTTTGGCTTCATGTTGGCAACAATATCATCTGGGCT 579
 Db 137 ACATTCACCTTCCTGATTTTGGCTTCATGTTGGCAACAATATCATCTGGGCT 78
 QY 580 GTCCTTCGTGACGAGATGTCATTCGACAAAGTTAACTGAGCTTGGAAGACTGACCTA 639
 Db 77 GTCCTTCGTGACGAGATGTCATTCGACAAAGTTAACTGAGCTTGGAAGACTGACCTA 18
 QY 640 CAAAGCA 647
 Db 17 AAGAGCA 10
 RESULT 14
 AOS12968 530 bp DNA linear GSS 05-MAY-1999
 LOCUS HS_5139_A2.G04.T7A RPCI-11 Human Male BAC Library Homo sapiens
 DEFINITION genomic clone Plate=715 Col=8 Row=M, genomic survey sequence.
 ACCESSION AOS12968
 VERSION AOS12968.1 GI:4745259
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 QY 1 (bases 1 to 530)
 Db Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 QY Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 JOURNAL 99380589
 MEDLINE 10449764
 PUBMED
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Libraries are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.bufileo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.bufileo.edu/ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end Web Server:
 http://www.htsc.washington.edu
 Plate: 715 row: M column: 8
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 530.
 Location/Qualifiers
 1..530
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=715 Col=8 Row=M"
 /sex="male"
 /clone_1db="RPCI-11 Human Male BAC Library"
 /note="Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"
 BASE COUNT 120 a 91 c 92 g 222 t 5 others

ORIGIN
 Query Match 24.8%; Score 461; DB 28; Length 530;
 Best Local Similarity 93.7%; Pred. No. 1.5e-62;
 Matches 479; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 678 ATTTCAGAAATTCAGAAAAATTCCTGAGTTCTTTGGCTTATTCATGCTTTAAAAA 737
 Db 12 ATTGAGAGCTGTGATTCATATATTCCTGCTTCCTGCTTATTCATGCTTTAAAAA 71
 QY 738 CTTCAGTATCTCTTCTGTTGACCTGCCCCATATCTTCCTGAGATTTCTGTTTGTGC 797
 Db 72 CTTCAGTATCTCTTCTGTTGACCTGCCCCATATCTTCCTGAGATTTCTGTTTGTGC 131
 QY 798 TTTTGTAGTAAATGCTGGAATTTGATTTGGGATCGGAGTATGATGAGCTGCT 857
 Db 132 TTTTGTAGTAAATGCTGGAATTTGATTTGGGATCGGAGTATGATGAGCTGCT 191
 QY 858 TCATTTTCCTCAACTATCTACTCTTTTTCATTTACTCTCTTTTCTTCTCTCATCT 917
 Db 192 TCATTTTCCTCAACTATCTACTCTTTTTCATTTACTCTCTTTTCTTCTCTCATCT 251
 QY 918 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTCTAGTTTGGCAACATGGAATTCGTT 977
 Db 252 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTCTAGTTTGGCAACATGGAATTCGTT 311
 QY 978 TTTTGTAGTAAATGCTGGAATTTGATTTGGGATCGGAGTATGATGAGCTGCT 1037
 Db 312 TTTTGTAGTAAATGCTGGAATTTGATTTGGGATCGGAGTATGATGAGCTGCT 371
 QY 1038 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTCTAGTTTGGCAACATGGAATTCGTT 1097
 Db 372 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTCTAGTTTGGCAACATGGAATTCGTT 431
 QY 1098 TGCATTTCTGAATATTTTGTAGTCCAGCCTATATATTTGCTGTTGAGTATAGCTGA 1157
 Db 432 TGCATTTCTGAATATTTTGTAGTCCAGCCTATATATTTGCTGTTGAGTATAGCTGA 491
 QY 1158 CTCATTTGAATCAAGCCAAATTTTGTGAAT 1188
 Db 492 CTCATTTGAATCAAGCCAAATTTTGTGAAT 522
 RESULT 15
 BY740968 724 bp mRNA linear EST 31-DEC-2002
 LOCUS BY740968 RIKEN full-length enriched, melanocyte Mus musculus cDNA
 DEFINITION Clone G270054A15 5', mRNA sequence.
 ACCESSION BY740968
 VERSION BY740968.1 GI:27165555
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 QY 1 (bases 1 to 724)
 Db Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
 Nakai,I., Osato,N., Saito,R., Suzuki,H., Yamanka,I., Kiyosawa,H.,
 Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,
 Gajbordi,T., Baldarelli,R., Hill,D.P., Bull,C., Hume,D.A.,
 Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
 Betzel,K.W., Blake,J.A., Bradt,D., Brusci,V., Chochia,C., Corbani,
 L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,
 A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.,
 Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
 Jarvis,E.D., Kanai,A., Kawai,H., Kawasawa,Y., Kedzierski,R.M.,
 King,B.L., Konagaya,A., Kurochik,I.V., Lee,Y., Lenhard,B., Lyons,
 P.A., Maglott,D.R., Malais,L., Marchionni,L., McKenzie,L., Miki,
 H., Nagashima,T., Numa,K., Okido,T., Pavan,W.J., Perlea,G.,
 Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
 Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,
 B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,
 M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 169 seconds
(without alignments)
4849.991 Million cell updates/sec

Title: US-10-000-151B-1
Perfect score: 1857
Sequence: 1 atgcgcacgactagagaggtta.....aagtaatatgncncaaa 1857

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/prodata/2/1na/5B_COMB.seq:*
3: /cgn2_6/prodata/2/1na/6A_COMB.seq:*
4: /cgn2_6/prodata/2/1na/6B_COMB.seq:*
5: /cgn2_6/prodata/2/1na/PCUTUS_COMB.seq:*
6: /cgn2_6/prodata/2/1na/backfilsl1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	502	27.0	539	3	US-09-337-171-5
2	73.4	4.0	7218	1	US-08-232-463-14
3	47.4	2.6	19124	2	US-08-487-826B-13
4	45.4	2.4	5340	4	US-09-627-122-21
5	44.6	2.4	658	4	US-08-998-416-595
6	44.6	2.4	2223	1	US-08-257-073-4
7	44	2.4	1203	4	US-09-134-001C-75
8	43.4	2.3	376	2	US-08-623-906A-18
9	43.2	2.3	2555	2	US-08-693-457-3
10	43.2	2.3	2555	2	US-09-265-731-3
11	42.8	2.3	5923	3	US-09-064-922-3
12	42.4	2.3	240	1	US-08-628-417-6
13	41.8	2.3	1020	4	US-09-328-475C-43
14	41.8	2.3	1696	4	US-09-835-811-1
15	41.8	2.3	1976	4	US-09-920-759-10
16	41.6	2.3	319608	4	US-09-679-409-1
17	41.6	2.2	6755	4	US-08-931-999-4
18	41.6	2.2	152331	3	US-09-128-155-16
19	41.6	2.2	176373	3	US-09-128-155-17
20	41.4	2.2	116592	4	US-09-818-512-3
21	41.2	2.2	454	2	US-08-623-906A-6
22	41.2	2.2	1736	3	US-09-182-816-22
23	41.2	2.2	1736	3	US-09-182-816-24
24	41.2	2.2	1736	3	US-09-471-528-22
25	41.2	2.2	1736	3	US-09-471-528-24
26	41.2	2.2	1736	3	US-09-634-530-22
27	41.2	2.2	1736	3	US-09-634-530-24

C	28	41.2	2.2	2447	2	US-09-014-969-14	Sequence 14, Appl
C	29	41.2	2.2	1664976	4	US-08-916-421B-1	Sequence 1, Appl
C	30	41	2.2	3001	4	US-09-539-333D-220	Sequence 220, App
C	31	40.8	2.2	570	4	US-09-107-532A-3368	Sequence 3368, App
C	32	40.6	2.2	319608	4	US-09-539-333D-1	Sequence 1, Appl
C	33	40.4	2.2	6124	4	US-08-213-419B-3	Sequence 3, Appl
C	34	40.4	2.2	90541	4	US-09-759-359A-3	Sequence 3, Appl
C	35	40.2	2.2	16442	3	US-08-781-891-208	Sequence 208, App
C	36	40.2	2.2	16442	3	US-09-618-166-208	Sequence 208, App
C	37	40	2.2	1493	1	US-08-340-820-24	Sequence 24, Appl
C	38	40	2.2	1493	1	US-08-593-535-24	Sequence 24, Appl
C	39	40	2.2	4285	4	US-09-410-464-1	Sequence 1, Appl
C	40	39.8	2.1	1447	4	US-09-443-041A-27	Sequence 27, Appl
C	41	39.8	2.1	5852	1	US-07-867-106-2	Sequence 2, Appl
C	42	39.6	2.1	1798	4	US-09-797-906-1	Sequence 1, Appl
C	43	39.6	2.1	4467	1	US-08-565-907A-1	Sequence 1, Appl
C	44	39.6	2.1	4467	2	US-08-910-51B-1	Sequence 1, Appl
C	45	39.6	2.1	4467	2	US-08-909-425A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-337-171-5
; Sequence 5, Application US/09337171
; Patent No. 6262249
; GENERAL INFORMATION:
; APPLICANT: Kennedy, Giulia C
; TITLE OF INVENTION: PANCREATIC CANCER GENES
; FILE REFERENCE: 200130.454
; CURRENT APPLICATION NUMBER: US/09/337,171
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-337-171-5

Query Match 27.0%; Score 502; DB 3; Length 539;
Best Local Similarity 96.8%; Pred. No. 8.2e-123;
Matches 522; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY	1051	AATAGACATTATACCTTTCTATGCTGGAAGAGCTTTTCAAGATATGCAATTCGAAA	1110
DB	1	AATAGACATTATACCTTTCTATGCTGGAAGAGCTTTTCAAGATATGCAAGCTGTAAAA	60
QY	1111	TATTTGTTAGTTCAGCCTATATATTTGCTGGATGATACCTGATGGAATCA	1170
DB	61	TATTTGTTAGTTCAGCCTATATATTTGCTGGATGATACCTGATGGAATCA	120
QY	1171	AAGCATTATTTTGAATTAATGTTTTCATATGCTGTTCAATTTATAGTTCCTCAG	1230
DB	121	AATTCATATTTTGAATTAATGTTTTCATATGCTGTTCAATTTATAGTTCCTCAG	180
QY	1231	AAACCTGCAATTTGTTACTTCAATTTTACCTTAATGCTTAATAGCTTAACATACT	1290
DB	181	AAACCTGCAATTTGTTACTTCAATTTTACCTTAATGCTTAATAGCTTAACATACT	240
QY	1291	CTGCGCCCAACATCCAGACTGTTTGGAACGAGTTCATATGCTTAATTTTCATA	1350
DB	241	CTGCGCCCAACATCCAGACTGTTTGGAACGAGTTCATATGCTTAATTTTCATA	300
QY	1351	ACTTTTACATCTTCTGGAACAAGACTTTTCACTGGCCAAATATGACGACATTCAAAG	1410
DB	301	ACTTTTACATCTTCTGGAACAAGACTTTTCACTGGCCAAATATGACGACATTCAAAG	360
QY	1411	TTTATGTGTAATATGATGATATTTTGAACCTTAATAATGACCTTAATAA-TAGACAT	1469
DB	361	TTTATGTGTAATATGATGATATTTTGAACCTTAATAATGACCTTAATAATAGACAT	420

APPLICANT: Dvorak, Jan
APPLICANT: Halverson, Joy
TITLE OF INVENTION: Microsatellite Sequences for Canine
TITLE OF INVENTION: Genotyping
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/623,906A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-62282/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..79
OTHER INFORMATION: /note= "Nucleotides 1-79 are unique
FEATURE:
NAME/KEY: misc_feature
LOCATION: 80..229
OTHER INFORMATION: /note= "Nucleotides 80-229 are
OTHER INFORMATION: repeat sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 230..376
OTHER INFORMATION: /note= "Nucleotides 230-376 are
OTHER INFORMATION: unique flanking sequence"
US-08-623-906A-18
Query Match 2.3%; Score 43.4; DB 2; Length 376;
Best Local Similarity 52.5%; Pred. No. 0.048; Mismatches 86; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 853 TGCTTCATTTTCTCTCAACTATCTACTTTTTCATTACTCTCTTTTCTTTTCTTCT 912
DB 231 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 172
QY 913 CATCTCTCTCTCTCTAGCAAAATTAAGACTTTTCTTTCTTTAGTTTGAACAATGAAT 972
DB 171 TTCTTCTCTCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 112
QY 973 CTGTGTTTGGTGTACCTTAGCTCTGTCTGTTTGTAGTTGAANAATCACTATGCTCAT 1032
DB 111 CTTCCTTTCTTTCTTTCTTTCTTTCTTTCTTTGTATGTCCTTGGTGTGTTTATAGGCTT 52
QY 1033 A 1033
DB 51 A 51

RESULT 9
US-08-693-457-3
Sequence 3, Application US/08693457
Patent No. 5880330
GENERAL INFORMATION:
APPLICANT: Weigel et al., Detlef
TITLE OF INVENTION: SHOOT MERISTEM SPECIFIC PROMOTER
TITLE OF INVENTION: SEQUENCES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,457
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hallie, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/012001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2555 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-693-457-3
Query Match 2.3%; Score 43.2; DB 2; Length 2555;
Best Local Similarity 46.0%; Pred. No. 0.15; Mismatches 154; Indels 0; Gaps 0;
Matches 132; Conservative 1; Mismatches 154; Indels 0; Gaps 0;
QY 1377 TTTTCAGTGGCCAAATAGTCAAGCATTTCAAGGTTTATGCTAATATCAGTCAATTT 1436
DB 641 TTTTCCTTAACAAGAGTTCATTATCTTGAGCTTGGAGTTTGAAGAGTT 700
QY 1437 TGAACGTAAATAATGACCTTAATATAGACCATTTCTACAAGAACAAGTGAATAGGNG 1496
DB 701 TACATTTATTAAGAAAAAATATATCATATATATATATATATATATATATATATAT 760
QY 1497 AAAACATGAAATTTCTTTAGTGCAGTGTGCTCTCAATATACATTAAGTTTATTTT 1556
DB 761 TTCACTAATTAATGCAACAACATCAATAATAGGTAACATACATATATAGTTTGT 820
QY 1557 ATATATTTAAACATATATGAAGATTAATTAATGCAAGAACTGGGAAGCTTAGACTG 1616
DB 821 TTACTCTTTTAAAAAAGGGAATTAACCTAAGAGGCTATTTCTCTAATAATTTAGGAC 880
QY 1617 CTTCGAANGCCTGAATATAGGAATAATAAAMNCTTTCAGATATCT 1663
DB 881 ACTGATCATATGACAAAAAATGCAATATATCATATTAATTTTGT 927
RESULT 10
US-09-265-731-3
Sequence 3, Application US/09265731
Patent No. 6239329
GENERAL INFORMATION:
APPLICANT: Weigel et al., Detlef
TITLE OF INVENTION: SHOOT MERISTEM SPECIFIC PROMOTER
TITLE OF INVENTION: SEQUENCES

RESULT 15
US-09-920-759-10/c
; Sequence 10, Application US/09920759
; Patent No. 6537811
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker
; APPLICANT: Susan M. Freiler
; TITLE OF INVENTION: ANTISENSE MODULATION OF SAP-1 EXPRESSION
; FILE REFERENCE: RTS-0267
; CURRENT APPLICATION NUMBER: US/09/920,759
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 10
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (150)...(1367)
US-09-920-759-10

Query Match

2.3%; Score 41.8; DB 4; Length 1976;

Best Local Similarity 56.0%; Pred. No. 0.3; Mismatches 62; Indels 0; Gaps 0;

Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

OY 880 TTTTTCATTACTCTCTTTTTCCTTCATCTCTGCTCCTAGCAAAATTAG 939
||||| ||| | ||||| ||| | | | | | | | | | |
DB 1973 TTTTTCATTACTCTCTTTTTCCTTCATCTCTGCTCCTAGCAAAATTAG 1914
||||| ||| | ||||| ||| | | | | | | | | | |
OY 940 ACTTTCCTTCCTAGTTGGAACATGGAATCTGTTTGGTGTTACTAGTCTCT 999
| |||| | ||| | ||||| | | | | | | | | | | | |
DB 1913 ATTTTATTTCAATAGTTTTCCTCCCTACAAAGTGTTTGGTTACATGATGAGTTCTT 1854
| |||| | ||| | ||||| | | | | | | | | | | | |
OY 1000 GTGTTTTCAGTTGGAATTC 1020
| | | | | | | | | | | | | | | | | | | | | |
DB 1853 TGGTGTAAATTCGAGATTTC 1833
| | | | | | | | | | | | | | | | | | | | | |

Search completed: September 23, 2003, 11:05:01
Job time : 172 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 1530 Seconds
(without alignments)
2985.665 Million cell updates/sec

Title: US-10-000-151b-1
Perfect score: 1857
Sequence: 1 atggcgcagctagaggtta.....aagtaataatgncncnaaa 1857

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1660708 seqs, 122959015 residues
Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1810.6	97.5	1857	US-10-000-151b-1	Sequence 1, Appl 1
2	502	27.0	539	US-09-773-459-5	Sequence 5, Appl 1
3	485.8	26.2	910	US-09-814-353-21669	Sequence 21669, A
4	258.4	13.9	283	US-10-040-739-376	Sequence 376, App
5	71	3.7	14006	US-10-311-455-1692	Sequence 1692, App
6	68	3.7	14006	US-10-311-455-1692	Sequence 1692, App
7	65.8	3.5	16033	US-10-311-455-1377	Sequence 1377, App
8	64.8	3.5	9155	US-10-311-455-435	Sequence 435, App
9	64.2	3.5	9539	US-10-240-453-54	Sequence 54, Appl
10	62.6	3.4	4654	US-10-239-676-52	Sequence 52, Appl
11	62.2	3.3	7571	US-10-311-455-2196	Sequence 2196, App
12	62.2	3.3	17934	US-10-311-455-500	Sequence 500, App
13	62	3.3	17934	US-10-311-455-1692	Sequence 1692, App
14	61	3.3	6593	US-10-311-455-451	Sequence 451, App
15	61	3.3	6668	US-10-311-455-1669	Sequence 1669, App
16	60.8	3.3	6944	US-10-172-086-112	Sequence 112, App

17	60.8	3.3	18281	US-10-311-455-935	Sequence 935, App
18	60.6	3.3	16167	US-10-311-455-1056	Sequence 1056, App
19	60.6	3.3	16167	US-10-240-485-82	Sequence 82, Appl
20	60.4	3.3	6145	US-10-311-455-945	Sequence 945, App
21	59.8	3.2	6334	US-10-311-455-1186	Sequence 1186, App
22	59.4	3.2	525	US-10-198-846-1483	Sequence 1483, App
23	59.2	3.2	6485	US-10-311-455-1781	Sequence 1781, App
24	59	3.2	5750	US-10-311-455-1981	Sequence 1981, App
25	58.8	3.2	6070	US-10-311-455-1652	Sequence 1652, App
26	58.8	3.2	6070	US-10-240-485-132	Sequence 132, App
27	58.8	3.2	9664	US-10-311-455-71	Sequence 71, Appl
28	58.6	3.2	6301	US-10-311-455-26	Sequence 26, Appl
29	58.4	3.1	12007	US-10-311-455-690	Sequence 690, App
30	58.2	3.1	6485	US-10-311-455-1782	Sequence 1782, App
31	57.6	3.1	5236	US-10-311-455-323	Sequence 323, App
32	57.6	3.1	8413	US-10-240-485-49	Sequence 49, Appl
33	57.4	3.1	11996	US-10-240-485-45	Sequence 45, Appl
34	57.2	3.1	5823	US-10-240-453-256	Sequence 256, App
35	56.8	3.1	7657	US-10-311-455-1995	Sequence 1995, App
36	56.8	3.1	7657	US-10-239-676-185	Sequence 185, App
37	56.6	3.0	529	US-09-983-965-2109	Sequence 2109, App
38	56.6	3.0	5198	US-10-311-455-252	Sequence 252, App
39	56.4	3.0	446	US-09-960-352-3400	Sequence 3400, App
40	56.2	3.0	7445	US-10-311-455-824	Sequence 824, App
41	56.2	3.0	9539	US-10-240-453-53	Sequence 53, Appl
42	56.2	3.0	9539	US-10-239-676-51	Sequence 51, Appl
43	56.2	3.0	113515	US-10-311-455-2147	Sequence 2147, App
44	55.8	3.0	6161	US-10-311-455-595	Sequence 595, App
45	55.8	3.0	6960	US-10-311-455-2372	Sequence 2372, App

ALIGNMENTS

```
RESULT 1
US-10-000-151b-1
; Sequence 1, Application US/10000151b
; Publication No. US20030013136A1
; GENERAL INFORMATION:
; APPLICANT: Baisley, Jeffrey R.
; TITLE OF INVENTION: HUMAN KCR1 REGULATION OF HERG POTASSIUM CHANNEL BLOCK
; FILE REFERENCE: Vanderbilt Ref No. US20030013136A1 V00120; Attorney Docket No. US2
; CURRENT APPLICATION NUMBER: US/10/000,151B
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1857)
; OTHER INFORMATION: n is any nucleic acid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1422)
; OTHER INFORMATION: n is any nucleic acid
US-10-000-151b-1
```

```
Query Match 97.5% Score 1810.6; DB 14; Length 1857;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1857; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ATGGCGAGCTAGAGGCTTACTGTTCTCGCGGCTTGTAGCTGTACTTTTGTCTCC 60
Db 1 ATGGCGAGCTAGAGGCTTACTGTTCTCGCGGCTTGTAGCTGTACTTTTGTCTCC 60
Oy 61 TGCCTCTCTTCTCCGCTTACGCGGCGCTGAGAGCCCTTACATGACGAGATCTTC 120
Db 61 TGCCTCTCTTCTCCGCTTACGCGGCGCTGAGAGCCCTTACATGAGAGATCTTC 120
```

OY	121	CACCGCCCTGAGGGGACAGCGCTACTGATGAGGGGCAATTCCTCCCTTCCAGTGGGATCC	180
Db	121	CACTGCTTCAGGGGACAGCGCTACTGATGAGGGCAATTCCTCCCTTCCAGTGGGATCC	180
OY	181	ATGATTTACTACATTACCTGCGCTGTGACTGGTGTCAGTTGGAGAGTGTCAAACCTGCCATT	240
Db	181	ATGATTTACTACATTACCTGCGCTGTGACTGGTGTCAGTTGGAGAGTGTCAAACCTGCCATT	240
OY	241	TGGATCTTTGGATGAGTCTGAACATGTTGTCGTCCATCTGSGATGCTCAGATTTGTTAAT	300
Db	241	TGGATCTTTGGATGAGTCTGAACATGTTGTCGTCCATCTGSGATGCTCAGATTTGTTAAT	300
OY	301	CTTCTCCCTCAGTGTGGCACTCTCATTTACTATATTTGCTTTTCCCAAGGTACAAACC	360
Db	301	CTTCTCTCAGTGTGGCACTCTCATTTACTATATTTGCTTTTCCCAAGGTACAAACC	360
OY	361	AGAAACAAGCGTGCGCTCAAGTATCCAGAGAGCTGTGTCACATTAACACTAGCACTATTT	420
Db	361	AGAAACAAGCGTGCGCTCAAGTATCCAGAGAGCTGTGTCACATTAACACTAGCACTATTT	420
OY	421	CCAAACCTTAATTTTTTTAACTCTCCCTTATTATACAGAAAGCAGAGATCATTTTAACT	480
Db	421	CCAAACCTTAATTTTTTTAACTCTCTTATTATACAGAAAGCAGAGATCATTTTAACT	480
OY	481	CTTTTGGATTTGGATGAGTGTCTTTATGGAATCATAAACTTCAGCTCTCTTGATTT	540
Db	481	CTTTTGGATTTGGATGAGTGTCTTTATGGAATCATAAACTTCAGCTCTCTTGATTT	540
OY	541	TGTGGCTTCATGTTTGGCAAAACAAATTCATCTGGGCTGTCTCTGTGCAAGGGAATGTC	600
Db	541	TGTGGCTTCATGTTTGGCAAAACAAATTCATCTGGGCTGTCTCTGTGCAAGGGAATGTC	600
OY	601	ATTCGACAAAAGTTAACTGAGGCTGTGAAAACAGCTACAAAAGAAAGAGACAGACTT	660
Db	601	ATTCGACAAAAGTTAACTGAGGCTGTGAAAACAGCTACAAAAGAAAGAGACAGACTT	660
OY	661	CCACTCTTAAGAGACCACTTTCGAGAATTCGAAAAAATCTTTCAGTTCTTTGGCTTAT	720
Db	661	CCACTCTTAAGAGACCACTTTCGAGAATTCGAAAAAATCTTTCAGTTCTTTGGCTTAT	720
OY	721	TCCATGTCCCTTAAACAACTGAGATGCTTCTCGTTGACTGTGGCCCTACATCCTGTG	780
Db	721	TCCATGTCCCTTAAACAACTGAGATGCTTCTCGTTGACTGTGGCCCTACATCCTGTG	780
OY	781	GGATTTCTGTTTGTGCTTTTGTAGTAGTTAATGCTGGAAATGTTATTTGGCATCGAGT	840
Db	781	GGATTTCTGTTTGTGCTTTTGTAGTAGTTAATGCTGGAAATGTTATTTGGCATCGAGT	840
OY	841	AGTATGATGAAGCGCTGTCTTCACTTTTCCCTCAGACTATTTCACTTTTTCATTTACTCTCTT	900
Db	841	AGTATGATGAAGCGCTGTCTTCACTTTTCCCTCAGACTATTTCACTTTTTCATTTACTCTCTT	900
OY	901	TTTTCTTTTCCCTCAGCTGTCTCTCTCCTAGCAAAATTAAGACTTTCTTCTTAAGTTTGG	960
Db	901	TTTTCTTTTCCCTCAGCTGTCTCTCTCTCCTAGCAAAATTAAGACTTTCTTCTTAAGTTTGG	960
OY	961	AAACATGGAATTCGTTTTTGTGGTGAACCTTACTCTCTGTGTTTATAGTTTGGAAATTC	1020
Db	961	AAACATGGAATTCGTTTTTGTGGTGAACCTTACTCTCTGTGTTTATAGTTTGGAAATTC	1020
OY	1021	ACTTATGCTCATTAATACTGTCTAGACAGACAATAGACATTATTACTTTCTATGTGTGAAA	1080
Db	1021	ACTTATGCTCATTAATACTGTCTAGACAGACAATAGACATTATTACTTTCTATGTGTGAAA	1080
OY	1081	AGAGTTTTTTCAAGATATGCAATTCGAAATATTTGTAGTCCAGGCTATATATTGCT	1140
Db	1081	AGAGTTTTTTCAAGATATGCAATTCGAAATATTTGTAGTCCAGGCTATATATTGCT	1140
OY	1141	GGTTGGAGTATAGCTGACTCATTTAAATCAAAAGCAATTTTTTGGAAATTAATGTATTC	1200
Db	1141	GGTTGGAGTATAGCTGACTCATTTAAATCAAAAGCAATTTTTTGGAAATTAATGTATTC	1200
OY	1201	ATATGCTGTGTTCAATGTTATAGTTCCTCCAGAAACGCTGGAATTTGTTACTGTCATTTTA	1260

Db	1201	ATATCGTTGTTATTTAGTTCCTCGAAGAACTGCTGGAAATTTGTTACTTCATTTTA	1260
QY	1261	CCTTATGTCATTTATAGGCTTTACATATACCTGCGCCGCCACATCCAGACTGTTGTGAA	1320
Db	1261	CCTTATGTCATTTATAGGCTTTAAACATTAACCTGCGCCGCCACATCCAGACTGTTGTGAA	1320
QY	1321	CTGAGTTGCTATGCATATGTTTAATTTCAATACCTTTTACATCTTTCTGAAACAAGCTTTT	1380
Db	1321	CTGAGTTGCTATGCATATGTTTAATTTCAATACCTTTTACATCTTTCTGAAACAAGCTTTT	1380
QY	1381	CAGTGGCCAAATAGTCAGACATTCGAAGCTTTAGTGTATATCAGTATATTTTGA	1440
Db	1381	CAGTGGCCAAATAGTCAGACATTCGAAGCTTTAGTGTATATCAGTATATTTTGA	1440
QY	1441	CTGTAAATAGGACTTAATATATAGACCATTTCTACAAACAACTGAATAGGAGAA	1500
Db	1441	CTGTAAATAGGACTTAATATATAGACCATTTCTACAAACAACTGAATAGGAGAA	1500
QY	1501	CATGGAATTTCTTTTAGGTGACAGTGGTGGTCTTCAATTTACATTAAGTTTTTAAATAT	1560
Db	1501	CATGGAATTTCTTTTAGGTGACAGTGGTGGTCTTCAATTTACATTAAGTTTTTAAATAT	1560
QY	1561	ATTTTAAACATATGTAAGAAATTAAGTGGCAAGAACTGGGAAAGCTTAAAGACTGCTTC	1620
Db	1561	ATTTTAAACATATGTAAGAAATTAAGTGGCAAGAACTGGGAAAGCTTAAAGACTGCTTC	1620
QY	1621	AAAGCCCTGAATATAGGCAAAATTAANNTGTTNCAGATATCTCATATGCTCCTNNKNAT	1680
Db	1621	AAAGCCCTGAATATAGGCAAAATTAANNTGTTNCAGATATCTCATATGCTCCTNNKNAT	1680
QY	1681	GNTGGCCCTTNNCAANGCTTGGGAATGKTNTNMNTGNATTAAGTTNATTAANCTGGGNT	1740
Db	1681	GNTGGCCCTTNNCAANGCTTGGGAATGKTNTNMNTGNATTAAGTTNATTAANCTGGGNT	1740
QY	1741	GCTNNMATNACTTNNKNCCANCCWNNNNACNAGNNNTANNANTATTACAAAGCTC	1800
Db	1741	GCTNNMATNACTTNNKNCCANCCWNNNNACNAGNNNTANNANTATTATTCAAAGGCTC	1800
QY	1801	AGGTATATTTCTGACTGAAAGTGCCTCTNNACATTAAGTAAATATGNGCCCA	1857
Db	1801	AGGTATATTTCTGACTGAAAGTGCCTCTNNACATTAAGTAAATATGNGCCCA	1857
RESULT 2			
US-09-773-459-5			
: Sequence 5, Application US/09773459			
: Patent No. US20010016651A1			
: GENERAL INFORMATION:			
: APPLICANT: Kennedy, Giulia C			
: TITLE OF INVENTION: PANCREATIC CANCER GENES			
: FILE REFERENCE: 200130.454			
: CURRENT APPLICATION NUMBER: US/09/773,459			
: CURRENT FILING DATE: 2001-01-31			
: PRIOR APPLICATION NUMBER: US 09/337,171			
: PRIOR FILING DATE: 1999-06-21			
: NUMBER OF SEQ ID NOS: 15			
: SOFTWARE: FastSeq for Windows Version 3.0			
: SEQ ID NO 5			
: LENGTH: 539			
: TYPE: DNA			
: ORGANISM: Homo sapien			
US-09-773-459-5			
Query Match 27.0%; Score 502; DB 9; Length 539;			
Best Local Similarity 96.8%; Pred. No. 2.9e-111;			
Matches 522; Conservative 0; Mismatches 16; Indels 1; Gaps 1;			
QY	1051	AATGACATTAATCTTCTATGTGAGAAAGATTTTCCAAGATATGCAATTCGAA	1110
Db	1	AATGACATTAATCTTCTATGTGAGAAAGATTTTCCAAGATATGCAATTCGAA	60
QY	1111	TATTTGTTAGTCCAGCCTATATATTTGCTGTGGAGATAGCTGACATTTGAATCA	1170

|||||
Db 61 TATTGTAGTCCACCTATATATTTGCTGGTGGATATACCTACATTTGAATCA 120
QY 1171 AAGCCATTTTGGAAATTAATGTTTCATATGCTGTGTCATTTGTTAGTCTCAG 1230
Db 121 AAGTCAATTTTGGAAATTAATGTTTCATATGCTGTGTCATTTGTTAGTCTCAG 180
QY 1231 AAGTCAATTTTGGAAATTTGTTACTTATTTACCTATGTCATTTATGAGGCTTAACATAC 1290
Db 181 AAGTCAATTTGTTGTTGTTACTTATTTACCTATGTCATTTATGAGGCTTAACATAC 240
QY 1291 CTGCTCCACATCCAGACTTGTGTGAACCTGAGTGTCTANGCAAATTTGTTAATTCATA 1350
Db 241 CTGCTCCACATCCAGACTTGTGTGAACCTGAGTGTCTANGCAAATTTGTTAATTCATA 300
QY 1351 ACTTTTACATCTTTGTGAACAGACTTTTCACTGGCCAAATGATGAGCATTCAGAG 1410
Db 301 ACTTTTACATCTTTGTGAACAGACTTTTCACTGGCCAAATGATGAGCATTCAGAG 360
QY 1411 TTTATGTGTAATATGATGATTTTGAACCTGTAATAATGACATTAATTA-TAGACAT 1469
Db 361 TTTATGTGTAATATGATGATTTTGAACCTGTAATAATGACATTAATTAAGACAT 420
QY 1470 TTCTACAAAGACACTGTAATGAGGAAACATGCAATTTCTTTAGGTGACAGTGTGG 1529
Db 421 TTCTACAAAGACACTGTAATGAGGAAACATGCAATTTCTTTAGGTGACAGTGTGG 480
QY 1530 TCTTCAATTTACATGTTTCTTTTATATATATTTAAACATATGTAAGAAATTAAGTG 1588
Db 481 TCTTCAATTTACATGTTTCTTTTATATATATTTAAACATATGTAAGAAATTAAGTG 539

RESULT 3

US-09-814-353-21669/c
; Sequence 21669, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21669
; LENGTH: 910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 909, 910
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21669

Query Match 26.2%; Score 485.8; DB 12; Length 910;
Best Local Similarity 94.3%; Pred. No. 3.3e-107;
Matches 516; Conservative 0; Mismatches 27; Indels 4; Gaps 1;

QY 1 ATGGCCAGCTAGAGGGTTACTGTTCTCGGCCGCTTGAGCTGTACCTTTTGTAGTCC 60
Db 787 ATGGCCAGCTAGAGGGTTACTGTTCTCGGCCGCTTGAGCTGTACCTTTTGTAGTCC 728
QY 61 TGCCTCCTCTTCGCCCTTCACCGCGCGCTCGAGAGCCCTACATGACGAGATCTTC 120
Db 727 TGCCTCCTCTTCGCCCTTCACCGCGCGCTCGAGAGCCCTACATGACGAGATCTTC 668
QY 121 CACCTGCTTCAGCGCAGCGCTACTGTGAGGGCCATTTTCTCCCTTCCAGTGGATCCC 180
Db 667 CACCTGCTTCAGCGCAGCGCTACTGTGAGGGCCATTTTCTCCCTTCCAGTGGATCCC 608
QY 181 ATGATTAACATACCTAGCTGCTGCTACCGGTGTCAGTTGGAGTGGTCAACCTGGCAT 240
Db 607 ATGATTAACATACCTAGCTGCTGCTACCGGTGTCAGTTGGAGTGGTCAACCTGGCAT 548
QY 241 TGGATCTTGGATGCTGTGAACATGTTGCTGCTCCATTGGAGTGCATGATTTGTTAAT 300
Db 547 TGGATCTTGGATGCTGTGAACATGTTGCTGCTCCATTGGAGTGCATGATTTGTTAAT 488
QY 301 CTTCCTCTCAGTGTGGCACTTCTATTACTATATTTGCTTCCACAAGTACAAACC 360
Db 487 CTTCCTCTCAGTGTGGCACTTCTATTACTATATTTGCTTCCACAAGTACAAACC 428
QY 361 AGAACAAGGCTGCTCAAGTATCCAGAGAGTCTTGCAACATTAACTAGCAGTATTT 420
Db 427 AGAACAAGGCTGCTCAGAGTATCCAGAGAGTCTTGCAACATTAACTAGCAGTATTT 368
QY 421 CCAACACTTATTTTAACTCTTATATATATATATATATATATATATATATATATAT 480
Db 367 CCAACACTTATTTTAACTCTTATATATATATATATATATATATATATATATATAT 312
QY 481 CTTTTGCATTTGATGCTCTTATGGAATATATATATATATATATATATATATATAT 540
Db 311 ACTCTTCTGATGATGCTCTTATGGAATATATATATATATATATATATATATATAT 252
QY 541 TGTGCT 547
Db 251 TGTGCT 245

RESULT 4

US-10-040-739-376
; Sequence 376, Application US/10040739
; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavalite, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```
APPLICATION NUMBER: 09/036,520
FILING DATE: 03-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 376:
SEQUENCE CHARACTERISTICS:
LENGTH: 283 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 376:
US-10-040-739-376

Query Match      13.98; Score 258.4; DB 13; Length 283;
Best Local Similarity 97.88; Pred. No. 1.7e-52;
Matches 262; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      371 CTGCTCAAGTATCCAGAGAGTCTTGTCAACATTAACAGTACAGTATTTCCACACTT 430
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      12 CTGCTGAGTAAACAGAGAGTCTTGTCAACATTAACAGTACAGTATTTCCACACTT 71
QY      431 ATTTTAACTCTCTTATTAATACAGAGAGAGTCTTATTTTACTCTTTTGCAT 490
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      72 ATTTTAACTCTCTTATTAATACAGAGAGAGTCTTATTTTACTCTTTTGCAT 131
QY      491 ATTGAGTGTCTTATTAAGAAATCAATAAATCTGAGCTTCTGATTTTGTGCTCA 550
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      132 ATTGAGTGTCTTATTAAGAAATCAATAAATCTGAGCTTCTGATTTTGTGCTCA 191
QY      551 TGTTCGCAAAACAATATCATCTGGGCTGCTCTGTGCGAGGAATGTCATTCACAAA 610
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      192 TGTTCGCAAAACAATATCATCTGGGCTGCTCTGTGCGAGGAATGTCATTCACAAA 251
QY      611 AGTTAACTGAGGCTTGAGAAACTGAGCT 638
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      252 AGTTAACTGAGGCTTGAGAAACTGAGCT 279

RESULT 5
US-10-311-455-1670
Sequence 1670, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1670
LENGTH: 6668
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
NAME/KEY: unsure
LOCATION: 1936
OTHER INFORMATION: n is a or g or c or t
```

```
US-10-311-455-1670
Query Match      3.88; Score 71; DB 12; Length 6668;
Best Local Similarity 44.08; Pred. No. 2.6e-06;
Matches 302; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY      698 TTCTTCAGTCTTCTTTGGCTTATTCATGTCCTTTAAACCTGAGATGCTTTCTGTT 757
      || || || || || || || || || || || || || || || || || || || || ||
DB      2766 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2825
QY      758 TGAATGCGCCATACCTCTGAGATTCGTTTCTGCTTTTGTAGTTAAGTGA 817
      || || || || || || || || || || || || || || || || || || || || ||
DB      2826 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2885
QY      818 GAATTTGTAATGGGATCGAGATGATGATGAGGCTCTCTCATTTCTCAACTATCT 877
      || || || || || || || || || || || || || || || || || || || || ||
DB      2886 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2945
QY      878 ACTTTTTCATTTACTCTCTTTTCTTTTCTTTTCTCTCATTCCTGCTCTAGCAAAATTA 937
      || || || || || || || || || || || || || || || || || || || || ||
DB      2946 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3005
QY      938 AGACTTTTCTTCTTCTTATTTGGAACATGGAATCTGTTTGTGTTACCTTACTCT 997
      || || || || || || || || || || || || || || || || || || || || ||
DB      3006 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3065
QY      998 CTGTTTCTTATTTGGAATTCATTAATGCTCAATTAATCTGCTAGCAACAATAGAC 1057
      || || || || || || || || || || || || || || || || || || || || ||
DB      3066 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3125
QY      1058 ATTACTTCTTCTAGTGTGGAAGAATTTTTCAAAATATGCAATTCGAAATATTTGT 1117
      || || || || || || || || || || || || || || || || || || || || ||
DB      3126 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3185
QY      1118 TAGTTCAGGCTATATATTTGCTGTTGAGATAGCTGACTGTGAATTCAGAAAGCCAA 1177
      || || || || || || || || || || || || || || || || || || || || ||
DB      3186 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3245
QY      1178 TTTTCTGGAATTAATGTTTTCATTAATGCTGTTTCAATGTTAATGCTCAGAAAGTC 1237
      || || || || || || || || || || || || || || || || || || || || ||
DB      3246 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3305
QY      1238 TCGAATTCGTTACTTATTTTACCTTATGCTATTAATGCTTAACATACCTGCTC 1297
      || || || || || || || || || || || || || || || || || || || || ||
DB      3306 TTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3365
QY      1298 CCACATCCAGACTGTTTGAAGTGTGATGCAATGTTAATTTCACTAATCTTTT 1357
      || || || || || || || || || || || || || || || || || || || || ||
DB      3366 TTTGTAGTATTAAGGGAATAGAAAGTGTGAATTTTAAATATTTGAAATTTTGT 3425
QY      1358 ACATCTTCTGACACAGACTTTTCAGT 1384
      || || || || || || || || || || || || || || || || || || || || ||
DB      3426 TAATTTTCGAAAGTATGTTTGATTT 3452

RESULT 6
US-10-311-455-1931
Sequence 1931, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
```



```

: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 2424
: SEQ ID NO 1931
: LENGTH: 14006
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 8289, 8310, 8313
: OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1931

```

Query Match	3.7%	Score 68;	DB 12;	Length 14006;
Best Local Similarity	44.8%;	Pred. No. 2.2e-05;		
Matches 260; Conservative	0;	Mismatches 320;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 7
US-10-311-455-1377
: Sequence 1377, Application US/10311455
: Publication No. US20030143606A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of diseases
: TITLE OF INVENTION: cytosine methylation
: FILE REFERENCE: 5013 1014
: CURRENT APPLICATION NUMBER: US/10/311,455
: CURRENT FILING DATE: 2002-12-16
: PRIOR APPLICATION NUMBER: PCT/EP01/07537

```

```

: PRIOR FILING DATE: 2001-07-02
: PRIOR APPLICATION NUMBER: DE 10032529.7
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: DE 10043826.1
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 2424
: SEQ ID NO 1377
: LENGTH: 16033
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens
: OS-10-311-455-1377

```

Query Match	3.5%;	Score 65.8;	DB 12;	Length 16033;
Best Local Similarity	47.3%;	Pred. No. 8.2e-05;		
Matches 199;	Conservative 0;	Mismatches 222;	Indels 0;	Gaps 0;

QY	698	TTCTTCAGTTTCCTTTTGGCTATATCCAGTCCCTTTAAAAACCTTGAGTATGCTTCTCGTT	757
Db	3071	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3130
QY	758	TGACTTGCCCTACATCCCTCTCGGAGTTCTGTTTGTGCTTTTGTAGTATGATAGTGTG	817
Db	3131	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3190
QY	818	GAATTGTATTTGGGATCGGAGTACGATAGAACGCTGTCATTTTCCATATTTCCATATCT	877
Db	3191	TATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3250
QY	878	ACATTTTTTTCATTACTCTCTTTTTTTTCTTTTCCCTACATCCCTGCTCCTAGCAAAAATA	937
Db	3251	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3310
QY	938	AGACTTTTCTTCTTACTAGTTTGGAAACATGGAATTCGTGTTTGGGTACTACTAGTCT	997
Db	3311	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3370
QY	998	CTGTGTTTTAGTTTGGAAATTCACATATGCTCATTAATACTTGCTGACAGACAATAGAC	1057
Db	3371	TTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTATTTATTTATTTTTTTAGAACGAT	3430
QY	1058	ATTATACTTCTATGTGTGAAAGAGCTTTTTCAAAGATATGCAATTCGTAATAATATTTGT	1117
Db	3431	ATTAGGAGAAAGGTGGTTGAAAGAGGATTTAGAAATGTTATATAGTTTATGATATTTAT	3490
QY	1118	T 1118	
Db	3491	T 3491	

```

RESULT 8
US-10-311-455-435
Sequence 435, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIRENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Det
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIORITY APPLICATION NUMBER: PCT/EP01/07537
PRIORITY FILING DATE: 2001-07-02
PRIORITY APPLICATION NUMBER: DE 10032529.7
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: DE 10043826.1
PRIORITY FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 435
LENGTH: 9155

```



```

Db      464  GTTTTGAGAGTTGATTAATTTTTTTGGTTGATTAATTTGTTGTAATATTTTATNAGTAAG 523
OY      1099  GCAATTCGTGAANAATTTTGGTAGTCCAGCCATATATATTTCTGCTGGAGATNAGCTGAC 1158
Db      524  TTTTATTTATTTAGTATATGATTTTTTATTTATTTAGATTTTATTTGGTTTTTATTTAT 583
OY      1159  TCATTGAAATCAAGAGCCAAATTTTTTGGAAATTTAGTTTATGTTTTCATATAGCTGTGCATTGTT 1218
Db      584  TTTTATTTATTTTATATATATTTTATTTATTTATTAAGTGTATATTTTATTTATTTGTTTA 643
OY      1219  ATAGTTCCTGAGAAACGTGAGATTTTCGTACTTCATTTTACCTATATGTCATTTATAGG 1278
Db      644  GTTCTTATTTTGGGTTTTTTTGTAAATTTGTTGAGTTATTTTAAAGATGATATTTTGAAT 703
OY      1279  CTTAACATTAACCTCGCTCCACATCCAGACTGTTTGTGACAGTGTCTATGCAAT 1338
Db      704  TTTTGTATGATTAATTTATAGATTTTTTATTTTATTTTATTTTATTTTATTTATTTT 763
OY      1339  GTTATATTTCAATCTTTTACATCTTCTGACAGACCTTTTCAGTGGCCAAATATAGTCAG 1398
Db      764  GTTTTATTTGATGGGTTTTATTTTATTTTATTTTATGTTATTTAGTATTTGATTTGAT 823
OY      1399  GACATTTCAAGGTTTATGTGTGTAATATACAGTATATTTGAATCTTAAGATTAAGACTTAA 1458
Db      824  GTTTTATATATGTATATTTATGTTTAAATATATGTTTATTTTATTAATGAATTTGTTT 883
OY      1459  TAATATGACCATTTCTACAAAGAACACACTGATAGGNGGAAGCAATGCAATTTCTTTAGG 1518
Db      884  TGATATGGGAAAGATAGTTTATGATAGCTTTTGGGTTTTGTTTATTTAGGTTATTTAGG 943
OY      1519  TGCAGTGTGTCTTCAATTTACTTATGTTTATTTTAAATATATATTTTAAACATATGTAA 1578
Db      944  TGTATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTAGTAT 1003
OY      1579  AAATTAAGTC 1588
Db      1004  TAAATTTTTT 1013

RESULT 15
US-10-311-455-1669
: Sequence 1669, Application US/10311455
: Publication No. US20030143606A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENBROCK, Christian

```

```

? TITLE OF INVENTION:  Diagnosis of Diseases Associated with the Immune System by De
? TITLE OF INVENTION:  cytosine methylation
? FILE REFERENCE:  5013.1014
? CURRENT APPLICATION NUMBER:  US/10/311,455
? CURRENT FILING DATE:  2002-12-16
? PRIOR APPLICATION NUMBER:  PCT/EP01/07537
? PRIOR FILING DATE:  2001-07-02
? PRIOR APPLICATION NUMBER:  DE 10032529.7
? PRIOR FILING DATE:  2000-06-30
? PRIOR APPLICATION NUMBER:  DE 10043826.1
? PRIOR FILING DATE:  2000-09-01
? NUMBER OF SEQ ID NOS:  2424
? SEQ ID NO 1669
? LENGTH:  6668
? TYPE:  DNA
? ORGANISM:  Artificial Sequence
? FEATURE:
? OTHER INFORMATION:  chemically treated genomic DNA (Homo sapiens)
? FEATURE:
? NAME/KEY:  unsure
? LOCATION:  4733
? OTHER INFORMATION:  n is a or g or c or t
US-10-311-455-1669

Query Match      3.3%; Score 61; DB 12; Length 6668;
Best Local Similarity  49.8%; Pred. No. 0.00069;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 22785 Seconds
(Without alignments)
3334.175 Million cell updates/sec

Title: US-10-000-151b-1
Perfect score: 1857
Sequence: 1 atggcgacgactagaggttca.....aagtaataatgngcncnaa 1857

Scoring table: IDENTITY_NMC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_mu:*
21: em_or:*
22: em_ov:*
23: em_ov:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vtl:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1647.4	88.7	2390	9 HSA312278	AJ312278 Homo sapi
2	1645.4	88.6	2280	6 AX498580	AX498580 Sequence
3	1388.2	85.5	1820	6 BC033730	BC033730 Homo sapi
4	1359.8	73.2	1461	6 AX591660	AX591660 Sequence
5	1349.2	72.7	165301	9 AC117372	AC117372 Homo sapi
6	1349.2	72.7	179032	2 AC024000	AC024000 Homo sapi
7	1311.8	70.6	1489	6 AX591658	AX591658 Sequence
8	1298.4	69.9	176600	9 AC046130	AC046130 Homo sapi
9	1233.8	66.4	158647	2 AC024478	AC024478 Homo sapi
10	1217.2	65.5	2362	6 BD160531	BD160531 Primer fo
11	1217.2	65.5	2362	9 AK027657	AK027657 Homo sapi
12	1075	57.9	7628	10 RNU78090	U78090 Rattus norv
13	775.6	41.8	276372	2 AC125569	AC125569 Rattus no
14	746.8	40.2	152534	2 AC114560	AC114560 Mus muscu
15	725.2	39.1	807	6 BD148636	BD148636 Primer fo
16	534	28.8	208006	2 AC087877	AC087877 Mus muscu
17	343.2	18.5	469	6 AX079565	AX079565 Sequence
18	336.4	18.1	442	6 AX079646	AX079646 Sequence
19	258.4	13.9	283	6 BD060016	BD060016 Secreted
20	202.4	10.9	2043	3 AK112724	AK112724 Ciona int
21	174.6	9.4	856	6 BD148363	BD148363 Primer fo
22	174.6	9.4	2245	6 BD158186	BD158186 Primer fo
23	174.6	9.4	2245	9 AK023061	AK023061 Homo sapi
24	151.8	8.2	546	6 BD153420	BD153420 Primer fo
25	88.6	4.8	1541	3 AY128497	AY128497 Drosophil
26	86.8	4.7	120142	2 AC017890	AC017890 Drosophil
27	86.8	4.7	161175	3 AC010023	AC010023 Drosophil
28	86.8	4.7	281592	3 AE003546	AE003546 Drosophil
29	86.4	4.7	1524	3 DME431376	AJ431376 Drosophil
30	83.6	4.5	41327	3 CBR647J19	AC084676 Caenorhab
31	80.8	4.4	93791	2 AC138073	AC138073 Homo sapi
32	77.6	4.2	66993	2 AC138074	AC138074 Homo sapi
33	76	4.1	111882	2 AC115612	AC115612 Dictyoste
34	74.6	4.0	1576	8 BT002117	BT002117 Arabidops
35	74.6	4.0	1758	8 AY09540	AY09540 Arabidops
36	73.4	4.0	7218	6 I66494	I66494 Sequence 14
37	72.6	3.9	8750	3 AY028643	AY028643 Plasmodiu
38	72	3.9	6644	6 E23356	E23356 Virus vecto
39	72	3.9	7372	6 E23357	E23357 Virus vecto
40	72	3.9	7797	6 E23355	E23355 Virus vecto
41	72	3.9	7996	6 E23359	E23359 Virus vecto
42	71	3.8	6668	6 AX346599	AX346599 Sequence
43	71	3.8	289973	2 AC135678	AC135678 Rattus no
44	70.4	3.8	143585	2 AC013349	AC013349 Homo sapi
45	70.4	3.8	156550	2 AC015830	AC015830 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS HSA312278
DEFINITION Homo sapiens mRNA for alpha2-glucosyltransferase (ALG10 gene).
ACCESSION AJ312278
VERSION AJ312278.1 GI:14349124
KEYWORDS ALG10 gene; alpha2-glucosyltransferase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Burda, P. and Aebi, M.
TITLE The ALG10 locus of Saccharomyces cerevisiae encodes the alpha-1,2
glucosyltransferase of the endoplasmic reticulum.

QY 1261 CCTATGTCATTTATAGGCTTTACATTAATCTGCTCCACATCCAGACTTTGTTGAA 1320
|||||
Db 1354 CCTATGTCATTTATAGGCTTTACATTAATCTGCTCCACATCCAGACTTTGTTGAA 1413
QY 1321 CTGAGTGCATATGCAATGTTAAATTCATTAATCTTTTACATCTTTTGAACAAGACTTT 1380
|||||
Db 1414 CTGAGTGCATATGCAATGTTAAATTCATTAATCTTTTCACTTTTCTTCAACAAGACTTT 1473
QY 1381 CAGTGGCCAAATGATCAGGACATCAAGGTTTATGCTATATATCACTGATATTTTGA 1440
|||||
Db 1474 CAGTGGCCAAATGATCAGGACATCAAGGTTTATGCTATATATCACTGATATTTGAA 1533
QY 1441 CTGTAATAATGACTTTAATAA-TAGACCATTTCTACAAGAACAACACTGAATAGNGGAAA 1499
|||||
Db 1534 CTGTAATAATGACTTTAATAA-TAGACCATTTCTACAAGAACAACACTGAATAGNGGAAA 1593
QY 1500 ACATGGAATTTCTTTTACGTCAGTGTGTCTTCAAAATTAACATTTGTTTTTAATAA 1559
|||||
Db 1594 ACATGGAATTTCTTTTACGTCAGTGTGTCTTCAAAATTAACATTTGTTTTTAATAA 1653
QY 1560 TATTTTAACATATCTAGAAATTAAGTGAAGAAGAACTGGAAGCTTAAGACCTGCTT 1619
|||||
Db 1654 TATTTTAACATATCTAGAAATTAAGTGAAGAAGAACTGGAAGCTTAAGACCTGCTT 1713
QY 1620 CAANAGCCTGATATGGAATAATAANMNGTTTNCAGATATCTATATGCTCTCANNKNA 1679
|||||
Db 1714 CAANAGCCTGATATGGAATAATAANMNGTTTNCAGATATCTATATGCTCTCANNKNA 1773
QY 1680 TGTGTCGCCCTTMCANAGCTTGGAGATGTTNNMNTGATTAATTAANCTGGGNN 1739
|||||
Db 1774 TGTGTCGCCCTTMCANAGCTTGGAGATGTTNNMNTGATTAATTAANCTGGGNN 1833
QY 1740 TGTCTNNMNTNA 1751
|||||
Db 1834 TGTCTCAAAAAA 1845

RESULT 3
BC033730 1820 bp mRNA 1linear PRI 08-JUL-2002
LOCUS Homo sapiens, clone MGC:44885 IMAGE:5168319, mRNA, complete cds.
DEFINITION BC033730
ACCESSION BC033730.1 GI:21707441
VERSION MGC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1820)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Ahter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Maduro,Q.L., Mastello,C., Maskerl,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stanlipop,S., Thomas,P.J., Touchman,J.W.,

Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 68 Row: 1 Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14249543.
FEATURES
location/Qualifiers
1..1820
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:44885 IMAGE:5168319"
/tissue_type="Brain, adult medulla"
/clone_id="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
101..1522
/codon_start=1
/product="unknown (protein for MGC:44885)"
/protein_id="AAH33730.1"
/db_xref="gi:21707442"
CDS
1..1522
/translation="MAQLEGYFESALSTFVSCILFSAFSRRLRPYDEIRHLP
AQRCEGFSLSDWPMITLGLVLSIGTRPAIVFSEHVCSIGMRVNL
FSVNGFYLYLDFCKVOPRNKAASSIORNLTILAVFPTLYFENFLYTEAGSMFT
LEAYLMCLYGNHKTSAFECGFEMRQNLIIWAVGAGNVAQRLTEAMKELOKED
RLPRIGPFAERKILQFLAYSMFSKLSMLLITFVYILGFLFCFAFVNGGLV
GDRSHACHPOLPFLYFESFPLTFSPKILSPKILFVLRRLRFLFVNLVAV
FLVKKRTYAHKYLILANDRHYTFYVWKRYFORYEVKYLIVPAVYIFACMSIADSKS
IFNMIFPICLFVIVPOKLEFRFIFLPIVYIRLNIPLPRLSLICELSCYAAVNI
TFEFLNKTFPMNSODIORPM"
BASE COUNT 504 a 348 c 334 g 634 t
ORIGIN
Query Match 85.5%; Score 1588.2; DB 9; Length 1820;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1622; Conservative 1; Mismatches 44; Indels 1; Gaps 1;
QY 1 ATGGGCGAGCTAAGAGGTTACTGTTTCGCGCCGCTTGAGCTGATCTTTTATGTC 60
Db 101 ATGGGCGAGCTAAGAGGTTACTGTTTCGCGCCGCTTGAGCTGATCTTTTATGTC 160
QY 61 TGCCTCTCTCTCTCCGCTTACGCGGGCGCTGCGAGAGCCCTACATGAGACATCTTC 120
Db 161 TGCCTCTCTCTCTCCGCTTACGCGGGCGCTGCGAGAGCCCTACATGAGACATCTTC 220
QY 121 CACCTGCTCAGGCGGACCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGGATCC 180
Db 221 CACCTGCTCAGGCGGACCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGGATCC 280
QY 181 ATGATTACTAATTACCTGCGCTTGTACTGCTGTCAGTTGGAGTGTCACAACTGCCATT 240
Db 281 ATGATTACTAATTACCTGCGCTTGTACTGCTGTCAGTTGGAGTGTCACAACTGCCATT 340
QY 241 TGGATCTTGGATGGTGTGAACATGTGTCTGCTCCATTTGGAGTGTGATTTAT 300
Db 341 TGGATCTTGGATGGTGTGAACATGTGTCTGCTCCATTTGGAGTGTGATTTAT 400
QY 301 CTTCCTTCATGTTGGCACTTCTATTTACTATATTTGCTTTCCACAAGTACCAACC 360
Db 401 CTTCCTTCATGTTGGCACTTCTATTTACTATATTTGCTTTCCACAAGTACCAACC 460
QY 361 AGAACAAGGCTGCTCAATATCAGAGAGTCTGTCAACATTAACACTAGAGATATT 420
Db 461 AGAACAAGGCTGCTCAATATCAGAGAGTCTGTCAACATTAACACTAGAGATATT 520
QY 421 CCAACACTTATTTTAACTCTTATTAATACAGAGAGATCTATGTTTAACT 480
Db 521 CCAACACTTATTTTAACTCTTATTAATACAGAGAGATCTATGTTTAACT 580
QY 481 CTTTTCATATTTGATGTGCTTATATGGAATCATTAACCTTACGCTTCTTGATTT 540
|||||

Db 581 CTTTGGCGTATTGATGTGCTTTATGGAATCATATAAACCCTGACGCTTCTTGGAATTT 640
 QY 541 TGTGCTTCATGTTTGGCAACAATATCATCTGGGCTGTCTTCTGTGAGAGGAATGTC 600
 Db 641 TGTGCTTCATGTTTGGCAACAATATCATCTGGGCTGTCTTCTGTGAGAGGAATGTC 700
 QY 601 ATTGCAAAAAGTTAACTGAGCGCTTGAAAACTGACCTACAAAAAGAGAAGACAGACTT 660
 Db 701 ATTGCAAAAAGTTAACTGAGCGCTTGAAAACTGACCTACAAAAAGAGAAGACAGACTT 760
 QY 661 CCACCTATTAAGAGACCATTTGAGAAATTCAGAAAAATCTTGAGTTCTTTGGCTTAT 720
 Db 761 CCACCTATTAAGAGACCATTTGAGAAATTCAGAAAAATCTTGAGTTCTTTGGCTTAT 820
 QY 721 TCCATGCTCTTTAAAAAAGTATGATGCTTTTCTGTTGACTGGGCTTACATCTCTG 780
 Db 821 TCCATGCTCTTTAAAAAAGTATGATGCTTTTCTGTTGACTGGGCTTACATCTCTG 880
 QY 781 GGATTTCTGTTTGTGCTTTTGTAGTATGTTAATGTTGAATTTGCGATCGGAGT 840
 Db 881 GGATTTCTGTTTGTGCTTTTGTAGTATGTTAATGTTGAATTTGCGATCGGAGT 940
 QY 841 AGTCAGAAAGCCGTCTCATTTTCTCTCAACTATCTACTTTTTCATTACTCTCTT 900
 Db 941 AGTCAGAAAGCCGTCTCATTTTCTCTCAACTATCTACTTTTTCATTACTCTCTT 1000
 QY 901 TTTTCTTTTCCATCTCCGCTGCTCCAGCAAAATTAAGACTTTCTTCCATAGTTTG 960
 Db 1001 TTTTCTTTTCCATCTCCGCTGCTCCAGCAAAATTAAGACTTTCTTCCATAGTTTG 1060
 QY 961 AAACAGGAATTCGTGTTTGTGTTGTTTACCTTATGTCGTGTTTGTAGTTGAAATTC 1020
 Db 1061 AAACAGGAATTCGTGTTTGTGTTGTTTACCTTATGTCGTGTTTGTAGTTGAAATTC 1120
 QY 1021 ACTATGCTCATATAACTTGTGAGAGACAAATAGACATATATCTTCTATGTGTGAAA 1080
 Db 1121 ACTATGCTCATATAACTTGTGAGAGACAAATAGACATATATCTTCTATGTGTGAAA 1180
 QY 1081 AGAGTTTTCAAAGATATGCAATTCGAAATTTTGTAGTCCAGCCATATATTTGCT 1140
 Db 1181 AGAGTTTTCAAAGATATGCAATTCGAAATTTTGTAGTCCAGCCATATATTTGCT 1240
 QY 1141 GGTGAGATATGCTGACTCATTAATCAATCAAAAGCCATTTTGTGAATTAATGTTTTC 1200
 Db 1241 GGTGAGATATGCTGACTCATTAATCAATCAAAAGCCATTTTGTGAATTAATGTTTTC 1300
 QY 1201 ATATGCTGTGCTATGTTATATGTTCTCCAGAAAGCTGTGGAATTTGCTTCAATTTA 1260
 Db 1301 ATATGCTGTGCTATGTTATATGTTCTCCAGAAAGCTGTGGAATTTGCTTCAATTTA 1360
 QY 1261 CCTATGCTATTAATAGGCTTAACATATGCTGCTCCCAATCCAGACTTGTGTGAA 1320
 Db 1361 CCTATGCTATTAATAGGCTTAACATATGCTGCTCCCAATCCAGACTTGTGTGAA 1420
 QY 1321 CTGAGTGTGATGCAATTTGTTAATTTCAATCAATTTTACATCTTTGTGAACAAGACTT 1380
 Db 1421 CTGAGTGTGATGCAATTTGTTAATTTCAATCAATTTTACATCTTTGTGAACAAGACTT 1480
 QY 1381 CAGTGGCCAATATGTCAGACATTCAAAGSTTATGTGTAATATCATATTTTGAA 1440
 Db 1481 CAGTGGCCAATATGTCAGACATTCAAAGSTTATGTGTAATATCATATTTTGAA 1540
 QY 1441 CTGTAAAAATGAGCTTAATAA-TAGACATTTCTACAAAGAACACAGATAGGNGAAA 1499
 Db 1541 CTGTAAAAATGAGCTTAATAAATGACCATTTCTACAAAGAACACAGATAGGNGAAA 1600
 QY 1500 ACATGGAATTTCTTTAGGTGAGTGTGCTTTCAAAATTTACATTTAGTTTAAATATA 1559
 Db 1601 ACATGGAATTTCTTTAGGTGAGTGTGCTTTCAAAATTTACATTTAGTTTAAATATA 1660
 QY 1560 TATTTTAAACATATGTAGAAATTAAGTGCAAAAGACCTGGAAAGCTTATAGACCTGCTT 1619
 Db 1661 TATTTTAAACATATGTAGAAATTAAGTGCAAAAGACCTGGAAAGCTTATAGACCTGCTT 1720

QY 1620 CAANAGCCGTAATTAATGGAAAAATAAAMNGTTTNCAGATATCTCATATA 1667
 Db 1721 CAANAGCCGTAATTAATGGAAAAATAAAMNGTTTNCAGATATCTCATATA 1768
 RESULT 4
 AX591660
 LOCUS
 DEFINITION
 AX591660
 ACCESSION
 AX591660
 VERSION
 AX591660.1
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 1
 AUTHORS
 Guo, X., Li, L., Patturajan, M., Shinkets, R. A., Casman, S. J.,
 Malvanekar, U. M., Tchernev, V. T., Vernet, C. A., Spytek, K. A.,
 Shenoy, S. G., Alsobrook, J. P., Edinger, S., Peyman, J. A., Stone, D. J.,
 Ellerman, K., Gangolli, E. A., Boldog, F. L., Colman, S. D., Eisen, A. J.,
 Liu, X., Padigaru, M., Spaderna, S. K. and Zernhusen, B. D.
 Proteins and nucleic acids encoding same
 Patent: WO 0246409-A 21 13-JUN-2002;
 Curagen Corporation (US)
 TITLE
 JOURNAL
 FEATURES
 source
 1..1461
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 347 a 298 c 277 g 539 t
 ORIGIN
 Query Match 73.2%; Score 1359.8; DB 6; Length 1461;
 Best Local Similarity 97.7%; Pred. No. 4; 6e-382;
 Matches 1379; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 1 ATGGCGCAGCTAGAGGGTACTGTTTCTGCGGCCCTTGAGCTGTACCTTTTATGTCG 60
 Db 49 ATGGCGCAGCTAGAGGGTACTGTTTCTGCGGCCCTTGAGCTGTACCTTTTATGTCG 108
 QY 61 TGCCTCCTCTTCCGCGCTTCAGCCGCGCTCGAGAGCCCTACATGAGAGATCTTC 120
 Db 109 TGCCTCCTCTTCCGCGCTTCAGCCGCGCTCGAGAGCCCTACATGAGAGATCTTC 168
 QY 121 CACCTGCGTCAGCGCAGCGCTACTGTGAGGCGCATTTCTCCCTTCCAGTGGATCC 180
 Db 169 CACCTGCGTCAGCGCAGCGCTACTGTGAGGCGCATTTCTCCCTTCCAGTGGATCC 228
 QY 181 ATGATTAATCTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 Db 229 ATGATTAATCTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288
 QY 241 TGGATCTTTGGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 Db 289 TGGATCTTTGGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348
 QY 301 CTTCCTCTCAGTGTGGCAACTTCTATTAATTAATTTGCTTTTCCAGAGTCAACC 360
 Db 349 CTTCCTCTCAGTGTGGCAACTTCTATTAATTAATTTGCTTTTCCAGAGTCAACC 408
 QY 361 AGAAACAAGCGCTCCCAATATCCAGAGAGCTGTGCAACATTAACATAGAGATTTT 420
 Db 409 AGAAACAAGCGCTCCCAATATCCAGAGAGCTGTGCAACATTAACATAGAGATTTT 468
 QY 421 CCAACACTTATTTTATTTTAACTTCTTATTAATATACAGAAACAGATCTATGTTTACT 480
 Db 469 CCAACACTTATTTTATTTTAACTTCTTATTAATATACAGAAACAGATCTATGTTTACT 528
 QY 481 CTTTGTGCATATTTGATGTGCTTTATGAAATCATATAAATCTACAGCTTCTTGGATTT 540
 Db 529 CTTTGTGCATATTTGATGTGCTTTATGAAATCATATAAATCTACAGCTTCTTGGATTT 588

QY 541 TGTGGCTTCATGTTTGGCAAAATATCATCTGGGCTGTCTTGTGACGGAATGTC 600
 DB 589 TGTGGCTTCATGTTTGGCAAAATATCATCTGGGCTGTCTTGTGACGGAATGTC 648
 QY 601 ATTGCACAAAAGTTAACTGAGGCTTGGAAAACTGAGTACAAAAGGAAGACAGACTT 660
 DB 649 ATTGCACAAAAGTTAACTGAGGCTTGGAAAACTGAGTACAAAAGGAAGACAGACTT 708
 QY 661 CCACCTTTAAAGGACCATTTGACGAATTCGAAAAATTTCTGATTTCTTTGGCTTAT 720
 DB 709 CCACCTTTAAAGGACCATTTGACGAATTCGAAAAATTTCTGATTTCTTTGGCTTAT 768
 QY 721 TCCATGCTCTTTAAAAAATTGAGTATGCTTTTCTGTTGACTTGGCCCTACATCTCTG 780
 DB 769 TCCATGCTCTTTAAAAAATTGAGTATGCTTTTCTGTTGACTTGGCCCTACATCTCTG 828
 QY 781 GGATTTCTGTTTGTGCTTTGTAGTAAATGAGTGAATTTGTTATTTGGCGAGT 840
 DB 829 GGATTTCTGTTTGTGCTTTGTAGTAAATGAGTGAATTTGTTATTTGGCGAGT 888
 QY 841 AGTCATGAGGCTGTCTTCATTTTCTCACTATTTCTTCTTCTTCTTCTTCTTCTTCTT 900
 DB 889 AGTCATGAGGCTGTCTTCATTTTCTCACTATTTCTTCTTCTTCTTCTTCTTCTTCTT 948
 QY 901 TTTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
 DB 949 TTTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1008
 QY 961 AAACATGGAATTCGTTTGTGTTGTTACCTTACTGCTGTCTGTTTGAATTC 1020
 DB 1009 AAACATGGAATTCGTTTGTGTTGTTACCTTACTGCTGTCTGTTTGAATTC 1068
 QY 1021 ACTTATGCTATTAATTAATGCTGTCAGACAAATGACATTAATTAATTAATTAATTAAT 1080
 DB 1069 ACTTATGCTATTAATTAATGCTGTCAGACAAATGACATTAATTAATTAATTAATTAAT 1128
 QY 1081 AGAGTTTTCAAAGATATGCAATTCGAAATTTGTTAGTCCAGCCTATATATTTGCT 1140
 DB 1129 AGAGTTTTCAAAGATATGCAATTCGAAATTTGTTAGTCCAGCCTATATATTTGCT 1188
 QY 1141 GGTGGATGATGAGTCTGCTATGGAATCAAGGCAATTTTGGATTTAAATGTTTTC 1200
 DB 1189 GGTGGATGATGAGTCTGCTATGGAATCAAGGCAATTTTGGATTTAAATGTTTTC 1248
 QY 1201 ATATGCTGTTTCAATTTGTTATGTTCTCTCAAGAACTGCTGAATTCGTTACTTATTTTA 1260
 DB 1249 ATATGCTGTTTCAATTTGTTATGTTCTCTCAAGAACTGCTGAATTCGTTACTTATTTTA 1308
 QY 1261 CCTTATGCTATTTATGAGCTTAAACATACCTGCTCCACATCCAGACATTTGTTGAA 1320
 DB 1309 CCTTATGCTATTTATGAGCTTAAACATACCTGCTCCACATCCAGACATTTGTTGAA 1368
 QY 1321 CTGAGTGTGATGATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
 DB 1369 CTGAGTGTGATGATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1428
 QY 1381 CAGTGGCCAAATAGTACGACATTCAAAGCT 1411
 DB 1429 CAGTGGCCAAATAGTACGACATTCAAAGCT 1459

RESULT 5
 AC117372
 LOCUS Homo sapiens 12 BAC R611-266015 (Roswell Park Cancer Institute
 DEFINITION Human BAC Library) complete sequence.
 AC117372
 ACCESSION AC117372 GI:20303507
 VERSION AC117372.2
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 AUTHORS
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 165301)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshman,F.R., Allen,C.,
 Alshrocks,S.L., Amaralungue,H.C., Are,J.R., Ayale,M., Banks,T.,
 Barbara,J., Benton,M., Birmage,K., Blankenburg,K., Bonini,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Burch,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Daborne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-carroll,L., Deckerich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dim,H.H.,
 Doulnwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Embling,S.,
 Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
 Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Gara,N.,
 Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
 Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M.,
 Holloway,C., Hollins,B., Homel,F., Howard,S., Huber,J., Hulyk,S.,
 Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
 Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
 King,J., Korvah,J., Kovar,C., Katovic,J., Kureshi,A., Landry,N.,
 Leal,B., Lee,E., Lewis,L.C., Lewis,J., Li,J., Li,Z., Lichtarge,O.,
 Lien,C., Liu,J., Liu,W., Louisgeed,H., Lozada,R.J., Lu,X.,
 Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P.,
 Mondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
 McInerney,E., McLeod,M.P., Meador,M., Mei,G., Mercher,S.,
 Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
 Monabadi,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
 Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okunou,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shim,C., Shoshitari,N., Sisson,I.,
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Umanit,K., Vasquez,L., Vera,Y., Villalona,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S.,
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
 Weinstock,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 165301)
 Morley,K.C.
 Direct Submission
 Submitted (10-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 165301)
 Morley,K.C.
 Direct Submission
 Submitted (25-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 165301)
 Morley,K.C.
 Direct Submission
 Submitted (30-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Apr 25, 2002 this sequence version replaced gi:20127350.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the

Oy 726 GTCTTTAAAACTGAGTATGCTTTCTGTTGACTGGCCCTACATCCCTCTGGCAT 785
 |||||
 Db 138121 GTCTTTAAAACTGAGTATGCTTTCTGTTGACTGGCCCTACATCCCTCTGGCAT 138180
 |||||
 Oy 786 TCGTTTGGCTTTTGTAGTATGAGTATGAGGAATGTTATGAGCATGGAGTATGCA 845
 |||||
 Db 138181 TCGTTTGGCTTTTGTAGTATGAGTATGAGGAATGTTATGAGCATGGAGTATGCA 138240
 |||||
 Oy 846 TGAAGCTGTCTTCATTTCTCCTCAACTATCTTCTTTTTCATTTACTCTCTTTTTC 905
 |||||
 Db 138241 TGAAGCTGTCTTCATTTCTCCTCAACTATCTTCTTTTTCATTTACTCTCTTTTTC 138300
 |||||
 Oy 906 TTTTCCATCTCTCTGCTCTCTAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCT 965
 |||||
 Db 138301 TTTTCCATCTCTCTGCTCTCTAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCT 138360
 |||||
 Oy 966 TGAATTCCTGTTTGTGGTGTACCTAGTCTGCTGTTTGTATTTGGAATTCACCTTA 1025
 |||||
 Db 138361 TGAATTCCTGTTTGTGGTGTACCTAGTCTGCTGTTTGTATTTGGAATTCACCTTA 138420
 |||||
 Oy 1026 TGCTCATTAATTAATGCTAGACAGCAATAGACATTAATCTTCTATGTTGGAAGAAGT 1085
 |||||
 Db 138421 TGCTCATTAATTAATGCTAGACAGCAATAGACATTAATCTTCTATGTTGGAAGAAGT 138480
 |||||
 Oy 1086 TTTTCAAGATATGCAATTCGAAATATTTGTATGTTCCAGCCCTATATTTGCTGTTG 1145
 |||||
 Db 138481 TTTTCAAGATATGCAATTCGAAATATTTGTATGTTCCAGCCCTATATTTGCTGTTG 138540
 |||||
 Oy 1146 GAGTATGCTGACTCATTTGAATCAAGCCCAATTTTGTGAATTAATGTTTTCATATG 1205
 |||||
 Db 138541 GAGTATGCTGACTCATTTGAATCAAGCCCAATTTTGTGAATTAATGTTTTCATATG 138600
 |||||
 Oy 1206 CTGTCATATGTTATGTTCTCTCAGAAACGCTGGAATTCGTTACTTCAATTTTACCTTA 1265
 |||||
 Db 138601 CTGTCATATGTTATGTTCTCTCAGAAACGCTGGAATTCGTTACTTCAATTTTACCTTA 138660
 |||||
 Oy 1266 TGTATATTAATGCTTAACATTAATCTGCTCCCAATCCAGCTTGTGTTGGAACGAG 1325
 |||||
 Db 138661 TGTATATTAATGCTTAACATTAATCTGCTCCCAATCCAGCTTGTGTTGGAACGAG 138720
 |||||
 Oy 1326 TTTGCTATGCAATGTTTAAATTCATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1385
 |||||
 Db 138721 TTTGCTATGCAATGTTTAAATTCATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 138780
 |||||
 Oy 1386 GCCAATATGTCAGACATTCAAAGGTTTATGTTGTAATATCAGATATTTTGAACGTGA 1445
 |||||
 Db 138781 GCCAATATGTCAGACATTCAAAGGTTTATGTTGTAATATCAGATATTTTGAACGTGA 138840
 |||||
 Oy 1446 AAAATGCACTTAATAA-TAGGCCATTTCTCAAGAACAACCTGAATAGGNGGAAACATG 1504
 |||||
 Db 138841 AAAATGCACTTAATAA-TAGGCCATTTCTCAAGAACAACCTGAATAGGNGGAAACATG 138900
 |||||
 Oy 1505 GAATTTCTTTTGTAGTGTGCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1564
 |||||
 Db 138901 GAATTTCTTTTGTAGTGTGCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 138960
 |||||
 Oy 1565 TAAACATATGTAAGAAATTAAGTGGCAAGAAGCTGGGAAGCTTAAGACCTGCTCAAN 1624
 |||||
 Db 138961 TAAACATATGTAAGAAATTAAGTGGCAAGAAGCTGGGAAGCTTAAGACCTGCTCAAN 139020
 |||||
 Oy 1625 GCCGGAATTAATGGAAGAAATAAANNCTTNCAGATATCTCATTCGCTCANNKNAATGTTG 1684
 |||||
 Db 139021 GCCGGAATTAATGGAAGAAATAAANNCTTNCAGATATCTCATTCGCTCANNKNAATGTTG 139080
 |||||
 Oy 1685 GCCCTTMMCAAGCTGGGAATGTTNNMMNTGNTAAGTTAAATTAAGCTGGGNNATGCTN 1744
 |||||
 Db 139081 GCCCTTMMCAAGCTGGGAATGTTNNMMNTGNTAAGTTAAATTAAGCTGGGNNATGCTN 139140
 |||||
 Oy 1745 NMAATNACT 1753
 :|||
 Db 139141 CATATTACT 139149

RESULT 6
 AC024000/c
 LOCUS
 DEFINITION
 AC024000
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

AC024000 179032 bp DNA linear HTG 08-MAR-2000
 Homo sapiens chromosome 6 clone RP11-266015 map 6, WORKING DRAFT
 SEQUENCE: 35 unordered pieces.
 AC024000
 AC024000.2 GI:7210053
 HTG: HTGS_PHASE1: HTGS_DRAFT.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 179032)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 6, clone RP11-266015
 Unpublished
 2 (bases 1 to 179032)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bede,F., Boguslavskiy,L.,
 Boukhalter,B., Brown,A., Burkett,G., Campolano,A., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearrellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
 Fenesfor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Illey,I., Johnson,K., Jones,C., Kam,L., Karatas,A.,
 Klein,J., Landers,T., Largocque,K., Lehocsky,J., Levine,R.,
 Liu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N., McCarthy,M.,
 McEwan,P., McGurk,R., McKernan,K., McPheters,R., Meldrum,J.,
 Meneus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M.,
 Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
 Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Turrell,A.,
 Travers,M., Trifilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
 Wu,X., Wyman,D., Ye,W.J., Young,G., Zalnoun,J., Zimmer,A. and
 Zody,M.
 Direct Submission
 Submitted (20-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 8, 2000 this sequence version replaced gi:7008905.
 All repeats were identified using RepeatMasker:
 Smt, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center Project name: L7057
 Center clone name: 266_O_15
 ----- Summary Statistics
 Sequencing vector: M13: M7815: 100% of reads
 Chemistry: Dye-terminator Big Dye, 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 153613 bases at least Q40
 Consensus quality: 164872 bases at least Q30
 Consensus quality: 171071 bases at least Q20
 Insert size: 175632; sum-of-ctrls
 Quality coverage: 3.8 in Q20 bases; sum-of-ctrls

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 35 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1242: contig of 1242 bp in length

1243 1342: gap of 100 bp
1343 2622: contig of 1280 bp in length
2623 2722: gap of 100 bp
2723 3815: contig of 1093 bp in length
3816 3915: gap of 100 bp
3916 5846: contig of 1931 bp in length
5847 5946: gap of 100 bp
5947 8052: contig of 2106 bp in length
8053 8152: gap of 100 bp
8153 10003: contig of 1851 bp in length
10004 10103: gap of 100 bp
10104 13638: contig of 3535 bp in length
13639 13738: gap of 100 bp
13739 16505: contig of 2767 bp in length
16506 16605: gap of 100 bp
16606 19624: contig of 3019 bp in length
19625 19724: gap of 100 bp
19725 22156: contig of 2432 bp in length
22157 22256: gap of 100 bp
22257 24548: contig of 2292 bp in length
24549 24648: gap of 100 bp
24649 27929: contig of 3281 bp in length
27930 28029: gap of 100 bp
28030 30657: contig of 2628 bp in length
30658 30757: gap of 100 bp
30758 34014: contig of 3257 bp in length
34015 34114: gap of 100 bp
34115 35085: contig of 971 bp in length
35086 35185: gap of 100 bp
35186 39055: contig of 3870 bp in length
39056 39155: gap of 100 bp
39156 42847: contig of 3692 bp in length
42848 42947: gap of 100 bp
42949 47517: contig of 4570 bp in length
47518 47617: gap of 100 bp
47618 50774: contig of 3157 bp in length
50775 50874: gap of 100 bp
50875 54256: contig of 3382 bp in length
54257 54356: gap of 100 bp
54357 58391: contig of 4035 bp in length
58392 58491: gap of 100 bp
58492 63149: contig of 4658 bp in length
63150 63249: gap of 100 bp
63250 68843: contig of 5594 bp in length
68844 68943: gap of 100 bp
68944 74638: contig of 5695 bp in length
74639 74738: gap of 100 bp
74739 80805: contig of 6067 bp in length
80806 80905: gap of 100 bp
80906 86709: contig of 5804 bp in length
86710 86809: gap of 100 bp
86810 93296: contig of 6487 bp in length
93297 93396: gap of 100 bp
93397 100379: contig of 6983 bp in length
100380 100479: gap of 100 bp
100481 110208: contig of 9729 bp in length
110209 110308: gap of 100 bp
110309 118705: contig of 8397 bp in length
118706 118805: gap of 100 bp
118806 130465: contig of 11660 bp in length
130466 130565: gap of 100 bp
130566 141090: contig of 10525 bp in length
141091 141190: gap of 100 bp
141191 154350: contig of 13160 bp in length
154351 154450: gap of 100 bp
154451 166361: contig of 11911 bp in length
166362 179032: contig of 12571 bp in length.
179033 179032: contig of 12571 bp in length.

FEATURES
source

1. 179032
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

/chromosome="6"
/map="6"
/clone="RP11-266015"
/clone_1b="RP11 Human Male BAC"
1. 1242
/note="assembly_fragment"
1343. 2622
/note="assembly_fragment"
2723. 3815
/note="assembly_fragment"
3916. 5846
/note="assembly_fragment"
5947. 8052
/note="assembly_fragment"
8153. 10003
/note="assembly_fragment"
10104. 13638
/note="assembly_fragment"
13739. 16505
/note="assembly_fragment"
16606. 19624
/note="assembly_fragment"
19725. 22156
/note="assembly_fragment"
22257. 24548
/note="assembly_fragment"
24649. 27929
/note="assembly_fragment"
28030. 30657
/note="assembly_fragment"
30758. 34014
/note="assembly_fragment"
clone_end:T7
vector_side:right"
34115. 35085
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
35186. 39055
/note="assembly_fragment"
39156. 42847
/note="assembly_fragment"
42948. 47517
/note="assembly_fragment"
47618. 50774
/note="assembly_fragment"
50875. 54256
/note="assembly_fragment"
54357. 58391
/note="assembly_fragment"
58492. 63149
/note="assembly_fragment"
63250. 68843
/note="assembly_fragment"
68944. 74638
/note="assembly_fragment"
74739. 80805
/note="assembly_fragment"

Query Match 72.7%; Score 1349.2; DB 2; Length 179032;
Best Local Similarity 97.6%; Pred. No. 8.4e-280;
Matches 1356; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 366 CAAGGCTGGCTCAAGATCCAGAGAGTCTTGTATACATTAACCTGACGATATTTCCAC 425
|||||
Db 9587 CAAGGCTGGCTCAAGATCCAGAGAGTCTTGTATACATTAACCTGACGATATTTCCAC 9528
|||||

QY 426 ACTTATTTTAACTTCCTTATATACAGAGCAGATCATGTTTTCCTCTTT 485
|||||
Db 9527 ACTTATTTTAACTTCCTTATATACAGAGCAGATCATGTTTTCCTCTTT 9468
|||||

QY 486 TGCATATTTGATGTGCTTTATGGAATACATAAACTTCAGCCTTCCTGATTTGTGG 545
|||||
Db 9467 TGCATATTTGATGTGCTTTATGGAATACATAAACTTCAGCCTTCCTGATTTGTGG 9408
|||||

QY 546 CTTCATGTTTCGGCAACAAATATCATCTGGGCTGTCTTCTGTGAGGGAAATGTCATTC 605
|||||
Db 9407 CTTCATGTTTCGGCAACAAATATCATCTGGGCTGTCTTCTGTGAGGGAAATGTCATTC 9348
606 ACAAAGTAACTGAGGCTGGAAACAGACATCAAAAAGAGAGACAGATCCACC 665
|||||
Db 9347 ACAAAGTAACTGAGGCTGGAAACAGACATCAAAAAGAGAGACAGATCCACC 9288
666 TATTAAAGACCAATTCGAGATTCAGAAAAATTCCTGAGTTCTTTGGCTTATTCAT 725
|||||
Db 9287 TATTAAAGACCAATTCGAGATTCAGAAAAATTCCTGAGTTCTTTGGCTTATTCAT 9228
726 GTCTTTAAAACTGAGATTCATCTTTCTGTTGACCTGGCCATCATCTTCGGGAT 785
|||||
Db 9227 GTCTTTAAAACTGAGATTCATCTTTCTGTTGACCTGGCCATCATCTTCGGGAT 9168
786 TCTGTTTGTGCTTTGTAGTAGTAAATGAGGAAATGTTATGCGAGATGAGTCA 845
|||||
Db 9167 TCTGTTTGTGCTTTGTAGTAGTAAATGAGGAAATGTTATGCGAGATGAGTCA 9108
846 TGAAGCCGTCTCATTTTCTCAACATTCATCTTTTTCATTTTACCTCTTTTTC 905
|||||
Db 9107 TGAAGCCGTCTCATTTTCTCAACATTCATCTTTTTCATTTTACCTCTTTTTC 9048
906 TTTTCCATCTCTCTGTCCTAGCAAAATTAAGACTTTTCTTCTTCTTGTGAAACA 965
|||||
Db 9047 TTTTCCATCTCTCTGTCCTAGCAAAATTAAGACTTTTCTTCTTCTTGTGAAACA 9888
966 TGAATTCGTTTGTGAGTACCTGCTCTCTGTTTGTAGTTGGAATTTACCTTA 1025
|||||
Db 8967 TGAATTCGTTTGTGAGTACCTGCTCTCTGTTTGTAGTTGGAATTTACCTTA 8928
1026 TGCATTAATATCTGCTAGCAGACAAATATATCTTCTTCTTCTTGTGGAAGAGT 1085
|||||
Db 8927 TGCATTAATATCTGCTAGCAGACAAATATATCTTCTTCTTGTGGAAGAGT 8868
1086 TTTTCAAGATATGCAATTCGAAATATTTGTAGTTCCAGCCATATATTTGCTGTTG 1145
|||||
Db 8867 TTTTCAAGATATGCAATTCGAAATATTTGTAGTTCCAGCCATATATTTGCTGTTG 8808
1146 GAGATATGCTATCATTTGAATTCAGAACCAATTTTGGAAATTAAGTTTTCATATG 1205
|||||
Db 8807 GAGATATGCTATCATTTGAATTCAGAACCAATTTTGGAAATTAAGTTTTCATATG 8748
1206 CTTCATCTGTTTATGTTCTCAGAACTGCTGGAATTTGCTTCTTCTTCTTACCTTA 1265
|||||
Db 8747 CTTCATCTGTTTATGTTCTCAGAACTGCTGGAATTTGCTTCTTCTTCTTACCTTA 8688
1266 TGTCAATTAAGGCTTAACATTAATCTGCTCCCAATCCAGACTTGTGTAAGT 1325
|||||
Db 8687 TGTCAATTAAGGCTTAACATTAATCTGCTCCCAATCCAGACTTGTGTAAGT 8628
1326 TTTGCTATGCAATTTGTAATTTTCAATCTTTTACATCTTTTGAACAAGCTTTCAGT 1385
|||||
Db 8627 TTTGCTATGCAATTTGTAATTTTCAATCTTTTACATCTTTTGAACAAGCTTTCAGT 8568
1386 GCCAAATATGTCAGGACATTCAGAGGTTATGTTGTAATATCATGTAATTTTGAACCTTA 1445
|||||
Db 8567 GCCAAATATGTCAGGACATTCAGAGGTTATGTTGTAATATCATGTAATTTTGAACCTTA 8508
1446 AAAATGCACTTAATAA-TAGACATTTCTACAAAAGAACACTGAATAGGNGAAAAATG 1504
|||||
Db 8507 AAAATGCACTTAATAA-TAGACATTTCTACAAAAGAACACTGAATAGGNGAAAAATG 8448
1505 GAATTTCTTTAGTGCGAGTGCTGCTCAATTAATCAATTTTATATATATAT 1564
|||||
Db 8447 GAATTTCTTTAGTGCGAGTGCTGCTCAATTAATCAATTTTATATATATAT 8388
1565 TAAACATATGTAAGAAATTAAGTGCAAGAACTGGAAAGCTTAAGACCTGCTCAAN 1624
|||||
Db 8387 TAAACATATGTAAGAAATTAAGTGCAAGAACTGGAAAGCTTAAGACCTGCTCAAN 8328

QY 1625 GCGTGAATTAATGGCAAAATTAANNNGTTTNCAGATATCTCATTCGCTCANNKATGTC 1684
|||||
Db 8327 GCGTGAATTAATGGCAAAATTAANNNGTTTNCAGATATCTCATTCGCTCANNKATGTC 8268
QY 1685 GCGCTTNCANGCTTGGGAATGKTTNNMNTGNATTAAGTTNATTAACCTGGNNGTCTN 1744
|||||
Db 8267 GCGCTTNCANGCTTGGGAATGKTTNNMNTGNATTAAGTTNATTAACCTGGNNGTCTN 8208
QY 1745 NMMATNACT 1753
Db 8207 CATATTACT 8199
RESULT 7
AX591658 1489 bp DNA linear PAT 27-JAN-2003
LOCUS Sequence 19 from Patent WO0246409.
DEFINITION AX591658
ACCESSION AX591658
VERSION AX591658.1 GI:27950044
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 Guo,X., Li,L., Patturajan,M., Shimkets,R.A., Casman,S.J.,
Malayanar,U.M., Tchernyev,V.T., Verneet,C.A., Spytek,K.A.,
Shenoy,S.G., Alsbroock,J.P., Edinger,S., Peyman,J.A., Stone,D.J.,
Ellerman,K., Gangoli,E.A., Boldor,F.L., Colman,S.D., Eisen,A.J.,
Liu,X., Padigaru,M., Spaderna,S.K., and Zernhusen,B.D.
Proteins and nucleic acids encoding same
Patent: WO 0246409-A 19 13-JUN-2002;
Curagen Corporation (US)
Location/Qualifiers
FEATURES
source 1..1489
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 357 a 306 c 279 g 547 t
ORIGIN
Query Match 70.6%; Score 1311.8; DB 6; Length 1489;
Best Local Similarity 95.4%; Pred. No. 9.9e-272;
Matches 1374; Conservative 0; Mismatches 37; Indels 30; Gaps 1;
QY 1 ATGGGCGAGCTAGAGGGTTTACGTTTCGCGCGCGCTGAGCTGTACCTTTTATGTC 60
|||||
Db 49 ATGGGCGAGCTAGAGGGTTTACGTTTCGCGCGCGCTGAGCTGTACCTTTTATGTC 108
61 TGCCTCTCTCTCCGCTTACGCGCGGCTGCGAGAGCCCTACATGAGAGATCTTC 120
|||||
Db 109 TGCCTCTCTCTCCGCTTACGCGCGGCTGCGAGAGCCCTACATGAGAGATCTTC 168
QY 121 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCCATTTCTCCCTTTCCAGTGGATGCC 180
|||||
Db 169 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCCATTTCTCCCTTTCCAGTGGATGCC 228
181 ATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
|||||
Db 229 ATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 288
241 TGGATCTTTGGATGGTGAACATGTTGTCTCCATTTGGATGGATGGATGGATGGAT 300
|||||
Db 289 TGGATCTTTGGATGGTGAACATGTTGTCTCCATTTGGATGGATGGATGGATGGAT 348
301 CTTCCTCTCAGTGTGGCACTCTGATTTACTATATTTGCTTTTCCACAAGATACACC 360
|||||
Db 349 CTTCCTCTCAGTGTGGCACTCTGATTTACTATATTTGCTTTTCCACAAGATACACC 408
361 AGAAGACAG-----GCTGCTCAGATATACAGAGA 390
|||||
Db 409 AGAAGACAGATATGTTTCAAAATTAATTAATTAACAGTTTGTGCTGCTCAGATATACAGAGA 468

QY 391 GTCCTGCAACATTAACTAGACAGATATTTCCAACTATATTTTAACTTCCTTAT 450
 DB 469 GTCCTGCAACATTAACTAGACAGATATTTCCAACTATATTTTAACTTCCTTAT 528
 QY 451 TATACGAGACAGATATATTTTACTCTTTTGCATATTTGATGTCTTTATGA 510
 DB 529 TATACGAGACAGATATATTTTACTCTTTTGCATATTTGATGTCTTTATGA 588
 QY 511 AATCATAAACCTCAGCCTTCCTTGATTTTGTGGCTTCATGTTGGGCAACAATATC 570
 DB 589 AATCATAAACCTCAGCCTTCCTTGATTTTGTGGCTTCATGTTGGGCAACAATATC 648
 QY 571 ATCTGGGCTCTCTCTGTCGAGGAATGTCATTCACAAAGTAACTGAGGCTTGAAA 630
 DB 649 ATCTGGGCTCTCTCTGTCGAGGAATGTCATTCACAAAGTAACTGAGGCTTGAAA 708
 QY 631 ACTGAGCTACAAAGAAGAGACAGACTTCCACTATTAAGACATTTGAGAAATTC 690
 DB 709 ACTGAGCTACAAAGAAGAGACAGACTTCCACTATTAAGACATTTGAGAAATTC 768
 QY 691 AGAAAAATCTCAGTTCTTTGGCTTATTCAGTCCCTTAAAACTGATATGCTT 750
 DB 769 AGAAAAATCTCAGTTCTTTGGCTTATTCAGTCCCTTAAAACTGATATGCTT 828
 QY 751 TTCTGTTGACTTGGCCCTACATCTTCTGGGATTTCTGTTGCTTTGATAGT 810
 DB 829 TTCTGTTGACTTGGCCCTACATCTTCTGGGATTTCTGTTGCTTTGATAGT 888
 QY 811 AATGCTGGAATTTTATGCGCATTCGAGATGATGATGACCCGTCTTCATTTTCTCAA 870
 DB 889 AATGCTGGAATTTTATGCGCATTCGAGATGATGATGACCCGTCTTCATTTTCTCAA 948
 QY 871 CTAATCTACTTTTTCATTTACCTCTTTTCTTTTCTTCATCCCTGCTCCAGC 930
 DB 949 CTAATCTACTTTTTCATTTACCTCTTTTCTTTTCTTCATCCCTGCTCCAGC 1008
 QY 931 AAAATTAAGACTTTTCTTCTTCTAGTTGGAACATGGAATCTGTTTGTGGTACC 990
 DB 1009 AAAATTAAGACTTTTCTTCTTCTAGTTGGAACATGGAATCTGTTTGTGGTACC 1068
 QY 991 TTATGCTCTGCTTTTACTTTTGAATTCATCTATGCTATTAATCTTCTGACAGAC 1050
 DB 1069 TTATGCTCTGCTTTTACTTTTGAATTCATCTATGCTATTAATCTTCTGACAGAC 1128
 QY 1051 AATAGACATTAATCTTCTTCTTCTAGTTGGAACATGGAATCTGTTTGTGGTACC 1110
 DB 1129 AATAGACATTAATCTTCTTCTTCTAGTTGGAACATGGAATCTGTTTGTGGTACC 1188
 QY 1111 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGGAGTATGACTGACTTGAATCA 1170
 DB 1189 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGGAGTATGACTGACTTGAATCA 1248
 QY 1171 AAGCAATTTTGTGAATTAATGTTTTCATATGCTTGTTCATGTTATAGTCCAG 1230
 DB 1249 AAGCAATTTTGTGAATTAATGTTTTCATATGCTTGTTCATGTTATAGTCCAG 1308
 QY 1231 AACTGCTGGAATTTGCTTACTTCAATTTACCTTATGTCATTTATAGGCTTAACATPACT 1290
 DB 1309 AACTGCTGGAATTTGCTTACTTCAATTTACCTTATGTCATTTATAGGCTTAACATPACT 1368
 QY 1291 CTGCTCCACATCCAGACTTGTGTGAACTGAGTGTGATGCAATTTGTAATTCATA 1350
 DB 1369 CTGCTCCACATCCAGACTTGTGTGAACTGAGTGTGATGCAATTTGTAATTCATA 1428
 QY 1351 ACTTTTACATCTTTTGAACAAGACTTTTCACTGGCCAAATAGTCAGACATTCAAAGG 1410
 DB 1429 ACTTTTACATCTTTTGAACAAGACTTTTCACTGGCCAAATAGTCAGACATTCAAAGG 1488
 QY 1411 T 1411
 DB 1489 T 1489

RESULT 8
 AC046130
 LOCUS
 DEFINITION
 Homo sapiens 12 BAC RP11-847H18 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
 AC046130
 AC046130.25 GI:11496301
 HTG.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 176600)

REFERENCE

AUTHORS

Alstbrooks, S.L., Amaralunge, H.C., Ale, J.R., Banks, T.R., Allen, C., Benton, U., Bimaye, K., Blankenbush, K., Bonin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Dean, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emmerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Girelli, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homai, F., Howard, S., Huber, J., Hui, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, Y., King, L., Kovach, J., Kovar, C., Kratovic, J., Kuresh, A., Landry, N., Lee, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Louisedge, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mathley, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokweto, S., Ogih, M., Okwun, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojoubkan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shooshitari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swalek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vison, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, C., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Zorilla, R., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 176600)
 Worley, K.C.
 Direct Submission
 Submitted (13-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 176600)
 Worley, K.C.
 Direct Submission
 Submitted (01-DEC-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 176600)
 Worley, K.C.
 Direct Submission
 Submitted (18-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

COMMENT

Baylor Plaza, Houston, TX 77030, USA
On Dec 1, 2000 this sequence version replaced gi:11415072.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-helpebcm.tmc.edu

Phrap Value Range

CLONE LENGTH: This sequence does not necessarily represent the entire length of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 176600
Phrap values in estimate: 175849
Average error rate (BCM-Phrap estimate): 1.51359e-05
Fraction of Phrap values less than 40 : 0.0100029
Number of consensus changing edits: 3
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original-Context Edited-Context
48814 cctctgccta(n)ctgttagac cctctgccta(c)ctgttagac
48821 ctacttgta(n)gactccct ctacttgta(g)gactccct
158405 cctctctct(n)tatatatata cctctctct(a)tatatatata

----- Distribution of Quality < 40 Bases -----

10001	*						
9001	*						
8001	*						
7001	*						
6001	*						
5001	*						
4001	*						
3001	*						
2001	*						
1001	*						
01	*						
5	10	15	20	25	30	35	40

FEATURES
source

Version: 1.01 gxf.

repeat_region	607..992	/rpl_family="MUT1A2"
repeat_region	1220..1627	/rpl_family="MSTD"
repeat_region	3841..4334	/rpl_family="MUT1D"
repeat_region	5769..5973	/rpl_family="MUT1J"
repeat_region	7338..7711	/rpl_family="MER52A"
repeat_region	7887..7909	/rpl_family="AT_rich"
repeat_region	8070..8090	/rpl_family="AT_rich"
STS	8661..8881	/standard_name="G51086"
repeat_region	9908..10030	/db_xref="dbSTS:73714"
repeat_region	10613..10965	/rpl_family="MIR"
repeat_region	10966..11237	/rpl_family="THE1B"
repeat_region	13358..13766	/rpl_family="MSTA"
repeat_region	14250..14387	/rpl_family="MIR"
repeat_region	14392..14692	/rpl_family="MER46C"
repeat_region	15147..15147	/rpl_family="THE1B"
repeat_region	15378..15378	/rpl_family="MSTA"
repeat_region	15941..16369	/rpl_family="MSTB"
repeat_region	17727..17727	/rpl_family="L2"
repeat_region	18185..18226	/rpl_family="AT_rich"
repeat_region	18311..18837	/rpl_family="MER89"
repeat_region	19039..19547	/rpl_family="MER89"

Query Match 69.9%; Score 1298.4; DB 9; Length 176600;
Best Local Similarity 95.4%; Pred. No. 7.1e-269;
Matches 1325; Conservative 7; Mismatches 56; Indels 1; Gaps 1;

QY	366	CAAGCTGCTCCATCAAGTATCCAGAGAGCTTGTCAACATTAACTAGCAGTATTTCCAAAC	425
DB	87687	CAAGCTGCTCCATCAAGTATCCAGAGAGCTTGTCAACATTAACTAGCAGTATTTCCAAAC	87746
QY	426	ACTTATTTTCTTAACTCCCTTTATTAACAGAGCAGATCTATGTTTTTACCTTTT	485
DB	87747	ACTTATTTTCTTAACTCCCTTTATTAACAGAGCAGATCTATGTTTTTACCTTTT	87806
QY	486	TGCATATTGATGTCTTTATGGAATCAATAAATCTACAGCTTCTTGATTTTGTGG	545
DB	87807	TGCATATTGATGTCTTTATGGAATCAATAAATCTACAGCTTCTTGATTTTGTGG	87866

QY 546 CTTCATGTTTCGGCAACAATATCATCTGGGCTGTCCTCTGTCAGAGGAATGTCATTCG 605
 Db 87867 CTTCATGTTTCGGCAACAATATCATCTGGGCTGTCCTCTGTCAGAGGAATGTCATTCG 87926
 QY 606 ACAAAAGTTACTGAGGCTTGGAAAATGAGCTACAAAAGAGAGACAGACTTCACC 665
 Db 87927 ACAAAAGTTAACGAGGCTTGGAAAATGAGCTACAAAAGAGAGACAGACTTCACC 87986
 QY 666 TATTAAAGGACATTTGGCAAAATTCGAAAATTCCTTCAGTTCTTTGGCTTATCCAT 725
 Db 87987 TATTAAAGGACATTTGGCAAAATTCGAAAATTCCTTCAGTTCTTTGGCTTATCCAT 88046
 QY 726 GTTCCTTAAAACTGATGATGCTTTCTGTTGACTGTGGCCCTACATCTTCCTGGGAT 785
 Db 88047 GTTCCTTAAAACTGATGATGCTTTCTGTTGACTGTGGCCCTACATCTTCCTGGGAT 88106
 QY 786 TCTGTTTGTGCTTTTGTAGTACTTAATGCTGAATTTGTTATGGGATCGAGTAGTCA 845
 Db 88107 TCTGTTTGTGCTTTTGTAGTACTTAATGCTGAATTTGTTATGGGATCGAGTAGTCA 88166
 QY 846 TGAAGCTGTCATCTTCCTCCACATATCTACTTTTTCATTTACTCTCTTTTTC 905
 Db 88167 TGAAGCTGTCATCTTCCTCCACATATCTACTTTTTCATTTACTCTCTTTTTC 88226
 QY 906 TTTTCTCATCTCTGCTCTAGCAAAATTAAGACTTTTCTTCTTCTAGTTTGAACA 965
 Db 88227 TTTTCTCATCTCTGCTCTAGCAAAATTAAGACTTTTCTTCTTCTAGTTTGAACG 88286
 QY 966 TGAATTCGTGTTTGTGCTTACCTTAGTCTCTGTTTGTAGTTTGAATTCATCTTA 1025
 Db 88287 TGAATTCGTGTTTGTGCTTACCTTAGTCTCTGTTTGTAGTTTGAATTCATCTTA 88346
 QY 1026 TGTCTAAATATCTGCTAGCAGCAAAATACATTAATCTTCTATGCTGGAAAAGAT 1085
 Db 88347 TGTCTAAATATCTGCTAGCAGCAAAATACATTAATCTTCTATGCTGGAAAAGAT 88406
 QY 1086 TTTTCAAGATATGCAATTCGAAATATTTGTAGTCCAGCCATATATTTCTGCTGTTG 1145
 Db 88407 TTTTCAAGATATGCAATTCGAAATATTTGTAGTCCAGCCATATATTTCTGCTGTTG 88466
 QY 1146 GAGTATATGCTGACTCATTAATTAATCAAGCAATTTTGTGAATTAATGTTTTCATAG 1205
 Db 88467 GAGTATATGCTGACTCATTAATTAATCAAGCAATTTTGTGAATTAATGTTTTCATAG 88526
 QY 1206 CTGTCATGTTATATGTTCTCGAAGATCTGCGAATTCCTTACTTCTTACTTCTTA 1265
 Db 88527 CTGTCATGTTATATGTTCTCGAAGATCTGCGAATTCCTTACTTCTTACTTCTTA 88586
 QY 1266 TGTCAATTTATAGCTTAACATTAATCTGCTCCACATCCAGACTCATTTGTGACTGAG 1325
 Db 88587 TGTCAATTTATAGCTTAACATTAATCTGCTCCACATCCAGACTCATTTGTGACTGAG 88646
 QY 1326 TTGCTATGCAATTTGTTAATTTCAATCTTTTCACTGTAACAAGACTTTTCACTG 1385
 Db 88647 TTGCTATGCAATTTGTTAATTTCAATCTTTTCACTGTAACAAGACTTTTCACTG 88706
 QY 1386 GCCAAATAGCAGCACTTAAAGGTTATGCTGTAATATCAGATATTTTGAACGTGA 1445
 Db 88707 GCCAAATAGCAGCACTTAAAGGTTATGCTGTAATATCAGATATTTTGAACGTG 88766
 QY 1446 AATATGACTTAATTAATAGACATTTCTACAAAAGCACTGAAATAGNGAAGCAATG 1504
 Db 88767 AATATGACTTAATTAATAGACATTTCTACAAAAGCACTGAAATAGNGAAGCAATG 88826
 QY 1505 GAATTTCTTTAGTGCAGTGCTGCTTCAAAATTAATTAATTAATTAATTAATTAAT 1564
 Db 88827 GAATTTCTTTAGTGCAGTGCTGCTTCAAAATTAATTAATTAATTAATTAATTAAT 88886
 QY 1565 TAAACATATGTAAGAAATTAAGTGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1624
 Db 88887 TAAACATATGTAAGAAATTAAGTGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 88946
 QY 1625 GCCTGAATATATGGAATAAAMNNGTTTCAGATATCTCATCTCANNKATGNTG 1684

Db 88947 GCCTGAATAATGGAATAAATTAATTTGTTTACATATCATATCATCATATGTTG 89006
 QY 1685 GCCCTTMCNCAACTGCGGAATGCTTNNMTGTAATTAAGTTNATTAANCTGGGNNGTG 1744
 Db 89007 GCCCTTAAACCTGCGGAATGTTTGTATGTACAAAGTTATTAATAAAGCTGGATGCTT 89066
 QY 1745 NMTATNACT 1753
 Db 89067 CATATTACT 89075

RESULT 9
 AC024478/c 158647 bp DNA linear HTG 26-MAY-2000
 LOCUS Homo sapiens chromosome 12 clone RP11-776P23 map 12, WORKING DRAFT
 DEFINITION SEQUENCE, 31 unordered pieces.
 ACCESSION AC024478
 VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barne, N., Bede, F., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Burkett, G., Campolongo, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 Dearliano, K., Dewar, K., Diaz, J., S., Dodge, S., Domino, M., Doyle, M.,
 Fenesor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Plerrie, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Landers, T., Laroque, K., Lehocsky, J., Levine, R.,
 Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M.,
 McKean, P., McGurk, A., McKernan, K., McPherson, R., Melidrum, J.,
 Meneus, L., Mihova, T., Miranda, C., Mienna, V., Morrow, J., Naylor, J.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivar, T. M.,
 Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C.,
 Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Titrill, A.,
 Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
 Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and
 Zody, M.

Direct Submission
 Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 25, 2000 this sequence version replaced gi:7259764.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

Project Information
 Center project name: L7124
 Center clone name: 776_P-23
 Sequencing Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 135824 bases at least Q40
 Consensus quality: 147226 bases at least Q30
 Consensus quality: 152156 bases at least Q20
 Insert size: 139000; agarose-fp

Insert size: 155647; sum-of-contigs
Quality coverage: 3.6 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1512: contig of 1512 bp in length
* 1513 1612: gap of 100 bp
* 1613 3017: contig of 1405 bp in length
* 3018 3117: gap of 100 bp
* 3118 4746: contig of 1629 bp in length
* 4747 4846: gap of 100 bp
* 4847 6511: contig of 1665 bp in length
* 6512 6611: gap of 100 bp
* 6612 8674: contig of 2063 bp in length
* 8675 8774: gap of 100 bp
* 8775 11190: contig of 2416 bp in length
* 11191 11290: gap of 100 bp
* 11291 14133: contig of 2843 bp in length
* 14134 14233: gap of 100 bp
* 14234 16333: contig of 2100 bp in length
* 16334 16433: gap of 100 bp
* 16434 20417: contig of 3984 bp in length
* 20418 20517: gap of 100 bp
* 20518 23097: contig of 2580 bp in length
* 23098 23197: gap of 100 bp
* 23198 27526: contig of 4329 bp in length
* 27527 27626: gap of 100 bp
* 27627 32176: contig of 4550 bp in length
* 32177 32276: gap of 100 bp
* 32277 36605: contig of 4329 bp in length
* 36606 36705: gap of 100 bp
* 36706 41220: contig of 4515 bp in length
* 41221 41321: gap of 100 bp
* 41321 44907: contig of 3587 bp in length
* 44908 45007: gap of 100 bp
* 45008 48806: contig of 3759 bp in length
* 48807 48906: gap of 100 bp
* 48907 53219: contig of 4313 bp in length
* 53220 53319: gap of 100 bp
* 53320 58369: contig of 5050 bp in length
* 58370 58469: gap of 100 bp
* 58470 62150: contig of 3681 bp in length
* 62151 62250: gap of 100 bp
* 62251 66960: contig of 4710 bp in length
* 66961 67060: gap of 100 bp
* 67061 72166: contig of 5106 bp in length
* 72167 72266: gap of 100 bp
* 72267 77323: contig of 5057 bp in length
* 77324 77423: gap of 100 bp
* 77424 82913: contig of 5490 bp in length
* 82914 83013: gap of 100 bp
* 83014 89483: contig of 6470 bp in length
* 89484 89583: gap of 100 bp
* 89584 97274: contig of 7691 bp in length
* 97275 97374: gap of 100 bp
* 97375 104650: contig of 7276 bp in length
* 104651 104750: gap of 100 bp
* 104751 112482: contig of 7732 bp in length
* 112483 112582: gap of 100 bp
* 112583 121328: contig of 8746 bp in length
* 121329 121428: gap of 100 bp
* 121429 128177: contig of 6749 bp in length
* 128178 128277: gap of 100 bp
* 128278 141975: contig of 13698 bp in length
* 141976 143075: gap of 100 bp
* 143076 158647: contig of 16372 bp in length.

FEATURES
Source
Location/Qualifiers
1. .158647
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12"
/clone="RP11-776P23"
/clone_11b="RP11-11 Human Male BAC"
1. .1512
/note="assembly_fragment"
misc_feature
1613. .3017
/note="assembly_fragment"
misc_feature
3118. .4746
/note="assembly_fragment"
misc_feature
4847. .6511
/note="assembly_fragment"
misc_feature
6612. .8674
/note="assembly_fragment"
misc_feature
8775. .11190
/note="assembly_fragment"
clone_end:77
vector_side:left
11291. .14133
/note="assembly_fragment"
misc_feature
14234. .16333
/note="assembly_fragment"
misc_feature
16434. .20417
/note="assembly_fragment"
misc_feature
20518. .23097
/note="assembly_fragment"
misc_feature
23198. .27526
/note="assembly_fragment"
misc_feature
27627. .32176
/note="assembly_fragment"
misc_feature
32277. .36605
/note="assembly_fragment"
misc_feature
36706. .41220
/note="assembly_fragment"
misc_feature
41321. .44907
/note="assembly_fragment"
misc_feature
45008. .48806
/note="assembly_fragment"
misc_feature
48907. .53219
/note="assembly_fragment"
misc_feature
53320. .58369
/note="assembly_fragment"
misc_feature
58470. .62150
/note="assembly_fragment"
misc_feature
62251. .66960
/note="assembly_fragment"
misc_feature
67061. .72166
/note="assembly_fragment"
misc_feature
72267. .77323
/note="assembly_fragment"
misc_feature
77424. .82913
/note="assembly_fragment"
misc_feature
83014. .89483
/note="assembly_fragment"
misc_feature
89584. .97274
/note="assembly_fragment"
misc_feature
97375. .104650
/note="assembly_fragment"
misc_feature
104751. .112482
/note="assembly_fragment"
clone_end:SP6
vector_side:left
112583. .121328
misc_feature
Query Match 66.4%; Score 1233.8; DB 2; Length 158647;
Best Local Similarity 88.9%; Pred. No. 5.7e-255;
Matches 1299; Conservative 3; Mismatches 159; Indels 1; Gaps 1;

QY 366 CAAGGCTGCCTCAGATATCCAGAGAGCTCTTGCAACATTAACTAGACAGATATTTCCAA 425
DB 50198 CAAGGCTGCCTCAGATATCCAGAGAGCTCTTGCAACATTAACTAGACAGATATTTCCAA 50139
QY 426 ACTTATATTTTAACTCTTATATATACAGAGAGATCTATGTTTTTACTCTTTT 485
DB 50138 ACTTATATTTTAACTCTTATATATACAGAGAGATCTATGTTTTTACTCTTTT 50079
QY 486 TGCATATTTGATGCTCTTATGGAATATATAAACTTACAGCTTCTTGATTTTGTG 545
DB 50078 TGCATATTTGATGCTCTTATGGAATATATAAACTTACAGCTTCTTGATTTTGTG 50019
QY 546 CTTCATGTTTCGCAACAATATCATCTGGGCTCTTCTGTGCAAGGATGTCATTC 605
DB 50018 CTTCATGTTTCGCAACAATATCATCTGGGCTCTTCTGTGCAAGGATGTCATTC 49959
QY 606 ACAAAGTTACTGAGGCTTGGAATCTAGCTACAAAAGAGAAAGACAGACTTCCACC 665
DB 49958 ACAAAGTTACTGAGGCTTGGAATCTAGCTACAAAAGAGAAAGACAGACTTCCACC 49899
QY 666 TATTAAGGACATTGCAAGATTCAGAAATTCCTCAGTTCTTTGGCTATTCAT 725
DB 49898 TATTAAGGACATTGCAAGATTCAGAAATTCCTCAGTTCTTTGGCTATTCAT 49839
QY 726 GTCTTTTAAACTGAGATGCTTTCTGTTGACTTGCCCTACATCTCTGGAT 785
DB 49838 GTCTTTTAAACTGAGATGCTTTCTGTTGACTTGCCCTACATCTCTGGAT 49779
QY 786 TCTGTTTGTCTTTGTAGTATGTTGGAATTTGTTATTTGGCATTCGAGTAGTCA 845
DB 49778 TCTGTTTGTCTTTGTAGTATGTTGGAATTTGTTATTTGGCATTCGAGTAGTCA 49719
QY 846 TGAAGCTGCTCTTCTCAGCAACTATTTCTACTTTTCTACTACTCTCTTTTTC 905
DB 49718 TGAAGCTGCTCTTCTCAGCAACTATTTCTACTTTTCTACTACTCTCTTTTTC 49659
QY 906 TTTTCTCATCTCTGCTCTCAGCAAAATTAAGACTTTTCTCTTCTGTTTGAAGA 965
DB 49658 TTTTCTCATCTCTGCTCTCAGCAAAATTAAGACTTTTCTCTTCTGTTTGAAGA 49599
QY 966 TGAATTCGTTTGTGTTTACCTTACTCTCTGTTTGTGTTTGAATTCACCTA 1025
DB 49598 TGAATTCGTTTGTGTTTACCTTACTCTCTGTTTGTGTTTGAATTCACCTA 49539
QY 1026 TGCCTAATAATCTGCTAGAGACATGACATTTACTTTCTACTGTTGGAAGACT 1085
DB 49538 TGCCTAATAATCTGCTAGAGACATGACATTTACTTTCTACTGTTGGAAGACT 49479
QY 1086 TTTTCAAGATATGCAATCTGAAATTTGTTAGTTCCAGCTATATTTGCTGTG 1145
DB 49478 TTTTCAAGATATGCAATCTGAAATTTGTTAGTTCCAGCTATATTTGCTGTG 49419
QY 1146 GAGTATAGCTGACTATTAATCAAGCAATTTTGTGAATTAATGTTTTCATAG 1205
DB 49418 GAGTATAGCTGACTATTAATCAAGCAATTTTGTGAATTAATGTTTTCATAG 49359
QY 1206 CTGTTGATTTTATAGTTCTCAGAAACTGCTGGAATTTTCTACTTATTTTACTTA 1265
DB 49358 CTGTTGATTTTATAGTTCTCAGAAACTGCTGGAATTTTCTACTTATTTTACTTA 49299
QY 1266 TGTCAATTTATAGGCTTAACATTAATCTGCTCCACATCCAGACTGTTTGTGAATGAG 1325
DB 49298 TGTCAATTTATAGGCTTAACATTAATCTGCTCCACATCCAGACTGTTTGTGAATGAG 49239
QY 1326 TTGCTATGCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1385
DB 49238 TTGCTATGCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 49179
QY 1386 GCGAATATGCTAGGACATTCAGAAAGTTTGTGTTATTAATTAATTAATTAATTAAT 1445
DB 49178 GCGAATATGCTAGGACATTCAGAAAGTTTGTGTTATTAATTAATTAATTAATTAAT 49119
QY 1446 AAAATGACTTAATAA-TAGACCATTCTTACAAAGAACATGAATAGGNGAAGAACATG 1504

DB 4918 AAAATGACTTAATAA-TAGACCATTCTTACAAAGAACATGAATAGGNGAAGAACATG 49059
QY 1505 GAATTTCTTTTATAGTGCAGATGGTGTCTTCAATTAATTAATTAATTAATTAATTT 1564
DB 49058 GAATTTCTTTTATAGTGCAGATGGTGTCTTCAATTAATTAATTAATTAATTAATTT 48999
QY 1565 TAAACATATGTAAGAAATTAAGTGGCAAGAACTGGAAAGCTTAAGACTGCTCAAN 1624
DB 48998 TAAACATATGTAAGAAATTAAGTGGCAAGAACTGGAAAGCTTAAGACTGCTCAAN 48939
QY 1625 GCCTGATATAGGGAATAAANMGTTCACATATCTCATATCCCTCANNATGNTG 1684
DB 48938 GCCTGATATAGGGAATAAANMGTTCACATATCTCATATCCCTCANNATGNTG 48879
QY 1685 GCCCTTMCANAGCTTGGCAATGKTTNNMNTGATTAAGTTAATTAANCTGGGNTGCTN 1744
DB 48878 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 48819
QY 1745 NMATNACTTNNKNCANCMNNNNACATGNNNTANNANATTTTACAAGCTCAGT 1804
DB 48818 NNNNNNNNNNNNTTACCATCAGTATATTTGTTTCTTATTCATCTTCTGCTG 48759
QY 1805 GATATCTTGACTGAAAGTGC 1826
DB 48758 TAGAATATTAAGTATAAGGCT 48737

RESULT 10
BD160531.
LOCUS BD160531 2362 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD160531
VERSION BD160531.1 GI:27866289
KEYWORDS JP 2002191363-A/15374.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2362)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K., and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 15374 09-JUL-2002;
JOURNAL
TITLE
AUTHORS
REFERENCE
COMMENT
OS Homo sapiens (human)
PN JP 2002191363-A/15374
PD 09-JUL-2002
PE 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAZUO
SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUO OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68, C12P21/08, G06F1/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT CDS
1..2362
Location/Qualifiers
1..2362
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 687 a 398 c 413 g 864 t
ORIGIN
Query Match 65.5%; Score 1217.2; DB 6; Length 2362;
Best Local Similarity 95.0%; Pred. No. 2.2e-251;
Matches 1245; Conservative 7; Mismatches 58; Indels 1; Gaps 1;
QY 444 CCTTATTTATACGAAGACGATCTATGTTTTTACTCTTTTTCATATTTGATGTCGT 503

[illegible][illegible]

Matches 1226: Conservative 0; Mismatches 230; Indels 3; Gaps 1;				
QY	1	ATGGCAGAGCTAGAGGTTACTGTTTCTCGGCGCCCTTGAGCTGTACCTTTTATGATGTC	60	
Db	101	ATGGCAGAGCTAGAGGTTACTGTTTCTCGGCGCCCTTGAGCTGTACCTTTTATGATGTC	160	
QY	61	TGCGTCTCTTTCGCCCTTACGCCGGCGCTGCGAGAGCCCTACAGAGAGATCTTC	120	
Db	161	TGCGTCTCTTTCGCCCTTACGCCGGCGCTGCGAGAGCCCTACAGAGAGATCTTC	220	
QY	121	CACCTGCGTCCAGGCGAGCGCTACTGTGAGGGCCATTCTCCCTTCCAGAGGATGCC	180	
Db	221	CACCTGCGTCCAGGCGAGCGCTACTGTGAGGGCCCTTCTACAGAGAGAGATGCC	280	
QY	181	ATGATTAATACATTACCTGCGCTGTACCTGCTGTACAGTGGAGTGTCAAACCTGCCAT	240	
Db	281	ATGATTAATACATTACCTGCGCTGTACCTGCTGTACAGTGGAGTGTCAAACCTGCCAGC	340	
QY	241	TGATATCTTTGGATGCTGAACATGTTGTCTGTCCATTTGGAGTGTCAATTTGTTAT	300	
Db	341	TGATATCTTTGGATGCTGAACATGTTGTCTGTCCATTTGGAGTGTCAATTTGTTAT	400	
QY	301	CTTTCCTTCACTGTTGGCACTTCTATTTACTATATTTGTTTCCACAGATCAACCC	360	
Db	401	CTTTCCTTCACTGTTGGCACTTCTATTTACTATATTTGTTTCCACAGATCAACCC	460	
QY	361	AGAAACAAGGCTGCTCAAGTATCCAGAGAGCTGTTCACATTAACACTAGCAATATT	420	
Db	461	AGAAACAAGGCTGCTCAAGTATCCAGAGAGCTGTTCACATTAACACTAGCAATATT	520	
QY	421	CCAACACTTATTTTCTTCTCTTATATATACAGAGAGATCTATGTTTACT	480	
Db	521	CCAACACTTATTTTCTTCTCTTATATATACAGAGAGATCTATGTTTACT	580	
QY	481	CTTTTCATATTTGATGTGTCTTATGAAATCAATAAACCCTTCTGATGAT	540	
Db	581	CTTTTCATATTTGATGTGTCTTATGAAATCAATAAACCCTTCTGATGAT	640	
QY	541	TGTGGCTTCATGTTTGGCAAAATATATATATGAGGCTGTCTGTGACAGGGAATGC	600	
Db	641	TGTGGCTTCATGTTTGGCAAAATATATATATGAGGCTGTCTGTGACAGGGAATGC	700	
QY	601	ATTGCACAAAAGTTAATCAGAGCTTGGAAAACCTGAGCTAC---AAAAGAGAGACAGCA	657	
Db	701	ATTGCACAAAAGTTAATCAGAGCTTGGAAAACCTGAGCTAC---AAAAGAGAGACAGCA	760	
QY	658	CTTCCACCTATTAAGAAGCAATTTGACAAATTTCAATTTCTTTTCTTTTGGCT	717	
Db	761	CTTCCACCTATTAAGAAGCAATTTGACAAATTTCAATTTCTTTTCTTTTGGCT	820	
QY	718	TATTCATATGCTTTTAAACTGATATGCTTTTCTGTTGACTTGGCCCTACATCCTT	777	
Db	821	TATTCATATGCTTTTAAACTGATATGCTTTTCTGTTGACTTGGCCCTACATCCTT	880	
QY	778	CTGGGATTTCTGTTTGTCTTTGTATGATTAATGTTGAATTTGTTATTTGGCATGG	837	
Db	881	CTGGGATTTCTGTTTGTCTTTGTATGATTAATGTTGAATTTGTTATTTGGCATGG	940	
QY	838	AGTATGATGAAGCGCTGCTCATTTTCTCAACATTTCTTTTCTTTTCTTACTCTC	897	
Db	941	AGTATGATGAAGCGCTGCTCATTTTCTCAACATTTCTTTTCTTTTCTTACTCTC	1000	
QY	898	TTTTTTTCTTTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	957	
Db	1001	TTTTTTTCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1060	
QY	958	TGGAACAATGGAATTTCTTTTGTGTTTACCTTACCTGCTGTGTTTGTGTTTGA	1017	
Db	1061	TGGAACAATGGAATTTCTTTTGTGTTTACCTTACCTGCTGTGTTTGTGTTTGA	1120	
QY	1018	TTTCACTATGCTCAATTAATCTGTGTCAGACAAATACATTAATCTTCTATGTTGG	1077	
Db	1121	TTTCACTATGCTCAATTAATCTGTGTCAGACAAATAGGCATTTACACGTTTATGTGG	1180	

QY	1078	AAAAGAGTTTTTCAAGATATGCAATTTCTGAAAATTTTGTATGTTCCAGCCTATATATT	1137	
Db	1181	AAAAGAGATTTTCAAGAGATATGCAATTTCTGAAAATTTTGTATGTTCCAGCCTATATATT	1240	
QY	1138	GCTGTTGAGATATATGCTGACTCATTTGAAATCAAAAGCCAAATTTTGTGAATTTAATGTT	1197	
Db	1241	GCTGTTGAGGTTGATGCTGACTCATTTTAAATCAAAAGCTCAATTTTCTGGAATTTAATGTT	1300	
QY	1198	TTCAATATGCTTGTATGTTATGTTATGTTCTGAGAAAGCTGGAATTTCTGTTACTTCAAT	1257	
Db	1301	TTTGTATGCTTGTGTTGTTCTTCTACGTTCTCTGAGAAAGCTGGAATTTCTGTTACTTCAAT	1360	
QY	1258	TTACCTATGCTATTTATATGCTTATACATCTGCTCTCCACATTCAGACTGTTTGT	1317	
Db	1361	TTACCTATGCTATTTATATGCTTATACATCTGCTCTCCACATTCAGACTGTTTGT	1420	
QY	1318	GAACGAGTTCCTATGCAATTTGTAATTTCTAATCTTTTACATCTTTCTGACAAAGCT	1377	
Db	1421	GAACGAGTTCCTATGCAATTTTCTAATCTTTTACATCTTTCTGACAAAGCT	1480	
QY	1378	TTTCAGTGGCAAAATAGTCAGACATTCAGAGTTTATGTTGTAATATGATGATATT	1437	
Db	1481	TTTCAGTGGTGGGATAGTATGATGATGATGATGATGATGATGATGATGATGATGAT	1540	
QY	1438	GAACGTAAAAAATGACTT 1456		
Db	1541	GAACGTAAAAAATGACTT 1559		
RESULT 13				
AC125569				
Rattus norvegicus clone CH230-8E24, *** SEQUENCING IN PROGRESS ***				
9 unordered pieces.				
AC125569				
AC125569.2 GI:23269718				
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.				
KEYWORDS				
Rattus norvegicus (Norway rat)				
SOURCE				
ORGANISM				
Rattus				
1 (bases 1 to 276372)				
REFERENCE				
AUTHORS				
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalebechi, Y., Aoyagi, A., Ayodeji, M., Bace, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Celler, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Diaper, H., Dugan-Kocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, J., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, N., Guera, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenan, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,				

Morgan, M., Morris, K., Morris, S., Muniadasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, G., Olanrewaju, A., Pal, S., Parks, K., Pastermak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quito, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 276372)
Worley, K.C.
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 276372)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (03-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 23, 2002 this sequence version replaced gi:21628982.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence only reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GDFR
Center clone name: CH230-8E24
----- Summary Statistics -----
Assembly program: Phrap: version 0.990329
Consensus quality: 193368 bases at least Q40
Consensus quality: 199764 bases at least Q30
Consensus quality: 203624 bases at least Q20
Estimated insert size: 222869; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a "working draft" sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

247498 247597: gap of unknown length
* 247598 267848: contig of 20251 bp in length
* 267849 269569: gap of unknown length
* 269570 269669: contig of 1621 bp in length
* 269670 271536: gap of unknown length
* 271537 271636: gap of unknown length
* 271637 273538: contig of 1902 bp in length
* 273539 274638: gap of unknown length
* 274639 274757: contig of 1118 bp in length
* 274757 274857: gap of unknown length
* 274857 276372: contig of 1516 bp in length.
Location/Qualifiers
1. 276372
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-8E24"
1. 1111
/note="wgs_contig"
misc_feature
4781..6514
/note="wgs_contig"
misc_feature
259347..261207
/note="wgs_contig"
BASE COUNT 61072 a 41701 c 42174 g 61457 t 69968 others
ORIGIN
Query Match 41.8%; Score 775.6; DB 2; Length 276372;
Best Local Similarity 82.4%; Pred. No. 2.2e-156;
Matches 902; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
366 CAAGCGCTCCCAAGTATCCAGAGAGCTGTGTCAACATTTACATACATGATTTCCAAAC 425
||||| |
139847 CAAAGCTTCTTCAAGTATCCAGAGATCTGTCAACATTTACATACATGATTTCCAAAC 139906
426 ACTTTATTTTAACTTCTTATTTATACAGAAAGAGATCTATTTTAACTCTTTT 485
||||| |
139907 CCTTATTTTAACTTCTTATTTATACAGAAAGAGATCTTATTTTAACTCTTTT 139966
486 TGCATTTATGCTGCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 545
||||| |
139967 TGCATTTATGCTGCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 140026
546 CTTCATTTTTCGGAACAATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
||||| |
140027 CTTCATTTTTCGTAAGCAATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 140086
606 ACAAAGTTAACTGAGGCTTGGAAGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACT 662
||||| |
140087 ACAAAGTTAACTGAGGCTTGGAAGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACT 140146
663 ACCTATTTAAAGACATTTTGAGAAATGAGAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
||||| |
140147 CCGCGTTAAAGACCGCTGCGAGTCAAGAGTTCATGTTCTGCGATGCTC 140206
723 CATGCTTTTAAAGACTGATATGCTTCTGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
||||| |
140207 CATGCTTTTAAAGACTGATATGCTTCTGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 140266
783 ATTTCGTTTGTGCTTTTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
||||| |
140267 GTTGCACTTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 140326
843 TCATGAAGCGCTGCTTATTTCTGCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
||||| |
140327 TCATGAAGCGCTGCTTATTTCTGCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 140386
903 TTCTTTCCATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
||||| |
140387 TTCTTTCCATCTCTACTTCTGCGACCAAGTCAAGTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 140446
963 ACATGAATTTCTGTTTGTGTTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1022
||||| |

```

Db      140447  ACGTAGAGTCAGTCTCTGTGATTACTTGTAGTCTGGTATTTTGGTTGGAATTCAC 140506
OY      1023  TTATGCTCATTAATTAATCTTCTAGACAGACATTAAGACATTAATCTTCTATGTGTGGAAG 1082
Db      140507  TTAGTCATAGATTATTAACAGGACAGACATTAAGACATTAACGCTTTTATGTGTGGAAG 140566
OY      1083  AGTTTTCAGATATATGCAATTCGAAATTAATTTGTTAGTCCAGCCTATATATTTGCTGCG 1142
Db      140567  AGTTATTCAGACACATGAATATGTCACATTAATTTAGTTCAGCCTACATGTTTGTGCG 140626
OY      1143  TTGGAGTATAGTCATGACATTAAGAAATCAAGCAATTTTGGATTTAAAGTTTTCAT 1202
Db      140627  TTGGGCTTACCTGACCTTTTAAATTAAGCAATTTTCTGGAATTTAAAGTTTGTGCT 140686
OY      1203  ATGCTGTGTCATTTGTTATAGTTCCTCAAGAACTGCTGGAATTTCTTACTTCATTTTACC 1262
Db      140687  ATGCTGTGTCCTTCTTACGGTTCCTCAGAAACTACTAGAAATTCCTTACTTCATTTTACC 140746
OY      1263  TTATGTCATTTATAGGCTTAAACATTAACATCTGCTGCCACATCCAGACTGTTTGTGAAC 1322
Db      140747  ATACATTTATTTATAGCTCAACATGCTCTGCGACCAATCTAGACTGTTTGTGAAC 140806
OY      1323  GAGTTGCTATGCAATGTTGTAATTTCAATTTTACATCTTTTCAATCTTTTGAACAGACTTTTCA 1382
Db      140807  GGGTTGCTATGCAATTTGTAATTTTCAATTTTACATCTTTTCAATCTTTTGAACAGACTTTTCA 140866
OY      1383  GTGGCCAATAGTCAAGACATTTCAAGGTTTATGTGTAATTTAGTATATTTTGAAC 1442
Db      140867  GTGGCGGATAGTCAATGCAATCCAAAGTTTATGTGTAATTTAGTCAAGGCTATTTTGAAC 140926
OY      1443  GTAAATATGACTT 1456
Db      140927  CTAAGAGCCTATTT 140940

RESULT 14
AC114560/c 152534 bp DNA linear HTG 11-JUN-2003
LOCUS      Mus musculus clone RP23-54K6, *** SEQUENCING IN PROGRESS ***, 3
AC114560
AC114560.5 GI:31581666
KEYWORDS   HTG; HTGS_PHASE2; HTGS_FULFILL; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 152534)
            Birren, B., Nusbaum, C. and Lander, E.
            Mus musculus, clone RP23-54K6
            Unpublished
            2 (bases 1 to 152534)
            Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
            Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L.,
            Boukhgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
            Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
            Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, D.S., Dodge, S.,
            Fero, S., Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,
            Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
            Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
            Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
            Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
            Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
            McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
            Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
            Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
            Oliver, J., Peterson, K., Phunhhang, P., Pierre, N., Pollara, V.,
            Raymond, C., Rella, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
            Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
            Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
            Straus, K., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
            Topham, K., Travers, M., Trivis, N., Trigglio, J., Vassiliev, H.,
            Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

```

```

TITLE      Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
JOURNAL    Direct Submission
REFERENCE   Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            3 (bases 1 to 152534)
AUTHORS    Birren, B., Nusbaum, C., Lander, E., Abouelkell, A., Allen, N.,
            Anderson, S., Arachchi, H.M., Barina, N., Bastien, V., Bloom, T.,
            Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
            Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
            Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
            Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
            Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
            Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
            Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
            Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
            Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
            Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
            Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
            O'Neill, D., Oliver, J., Peterson, K., Phunhhang, P., Pierre, N.,
            Rachupka, A., Ramasamy, U., Raymond, C., Rella, R., Rise, C., Rogov, P.,
            Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
            Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
            Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
            Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
            Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
            Direct Submission
            Submitted (11-JUN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Jun 11, 2003 this sequence version replaced gi:30984966.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: MIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L23271
            Center clone name: 54_K_6
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            * 1 80469: contig of 80469 bp in length
            * 80470 80569: gap of 100 bp
            * 80570 90010: contig of 9441 bp in length
            * 90011 90110: gap of 100 bp
            * 90111 152534: contig of 62424 bp in length.
            Location/Qualifiers
            1..152534
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="RP23-54K6"
            /clone_11b="RPCT-23 Female Mouse BAC"
            BASE COUNT 44923 a 30812 c 30537 g 45587 t 675 others
            ORIGIN
            Query Match 40.2% Score 746.8 DB 2: Length 152534;
            Best Local Similarity 81.1% Pred. No. 3.4e-150;
            Matches 881; Conservative 0; Mismatches 202; Indels 3; Gaps 1;
            OY 366 CAAAGCTGCTCACTATCAAGAGAGTCTTGCAACATTAACACAGAGATTTTCAAC 425
            Db 52439 CAAAGCTTCTTCAAGTATCCAGAGATCTTGCAACATTAACCTTACAGAGTTCAC 52380

```

OY	426	ACCTTATTTTTTAACTCCCTTTATATACAGAAAGAGATCGATGTTTTTAACTCCTTTT	485
Db	52379	CCCTATTTTTTTTTTAACTCCCTTACTATATACAGAAAGCTGGGTGTATCTCACTCTTTT	523200
OY	486	TGCATATTTATGATGTCCTTTATGGAATCATATAACCTGACGCTTCCTTGATTTTGTGG	545
Db	52319	TGCTTATTTTGAATGTCCTTATGGAACCAATATAGACCTCGGCCCTTGCTTGATTTTGTGG	522660
OY	546	CTTCAATGTTTCGGCAACCAATATCATCTGGGCTGTCTTCTGTGCAGGGAATGTCATTGC	605
Db	52259	TTTTCATGTTTCGTGACACCAACATCACTGGGGGGGCTCTCTGTGCAGGACACCTCATATGC	522000
OY	606	ACAAAGCTTAACTAGAGCTTGGAAAACCTGAGTACAA --- ABAAGAAAGACAGCATTTGC	662
Db	52199	ACAAAGCTGAGTGAAGCTTGGAAAGATTGAGCTGCAGAAAGAAAGAAAGAGAGAGGCTTGC	521400
OY	663	ACCTATTTAAAGACCAATTTGCAGAAATTCAGAAAAATTCCTACGTTTCTTTGGCTTATTC	722
Db	52139	GCCCCACGAAGGAGCCGCTCTCGAGAGCTCAAAAGAGTGCATGTTCTCTGCTGTGACGC	520800
OY	723	CATCTCCCTTAAAAACCTGAGTATGCTTCTCTTTGACTTGGGCCCTACATCCTCTCGGG	782
Db	52079	CATCTCCCTTAAAGACCTGGAATGCTCTTCTGTGACCTGGCCCTACGTGCTCTGTGT	520200
OY	783	ATTTCCTGTTTGTGCTTTGTAGTAGTTAATGGTGAATTTGTTATGGGATCGAGTAG	842
Db	52019	GTTTGGCGTTCTTTGCCCTTTGTGGAGTTAATGGTGCAATGTTGTTGGTAGCCGAGTAG	519660
OY	843	TCATGAAGCGCTGCTCTTATTTTCTCTCAACTATTCCTACTTTTTCATTTACTCTCTTTT	902
Db	51959	TCACGAGGCGCTGCTCATTTTCCCTCAAGTGTCTATTTCTTCTCTTACGCCCTTTT	519000
OY	903	TTCTTTTCCATCTCTGCTGCTCTAGCAAAATTAAGACTTTTCTCTCTAGTTGGAA	962
Db	51899	CTCTTCTCCCTCACTACTTCTCTGACCAAGTCAAGACTTCTCTTACTAGCTTATTTGGAA	518400
OY	963	ACATGGAATTCGTGTTTGTGGGTTTACCTTAGTCTGTGCTTTTATGTTGAAATTCAC	1022
Db	51839	GCGTAGAGTCCAGTTCCTGTGGTTAGCTTACGTCGATCACTTTTGGTTGAAATTCAC	517800
OY	1023	TTTATGTCATAAATACTGCTGACAGACATATACATTAATCTTATGTTGGGAAAG	1082
Db	51779	TTATGTCATATAGTATTTACTGAGACCAATATAGCATTCACATTTTATGTTGGGAAAG	517200
OY	1083	AGTTTTTCAAGATATGCAATTCCTGAAATTTTGTAGTTCAGCCTATATATTTGCTGG	1142
Db	51719	AGATTTTCAAGAGACATGAATGTGCAAATTTTATAGTTCAGCCTACATTTTCTGG	516660
OY	1143	TTGCAATATAGCTGACTCATTTGAATCAAAAGCCAAATTTTGGAAATTAATGTTTCAAT	1202
Db	51659	TTGGGCAATATAGCTGACTCTTTAAAGGTAAAGTCAATTTCTGGAAATTAATGTTTGTGT	516000
OY	1203	ATGCTTGTTCATGTTTATGTTACTGTCAGAAAGCTGGGAATTCGTTACTCATTTTAC	1262
Db	51599	ATGCTTGTGTTCTTCTACACTTCTCTCAGAAACTACTAGATTCGTTACTCTCATTTTAC	515400
OY	1263	TTATGTCATTTATAGGCTTAACTAATCTGCTCCACATCCAGACTGTGTTGTAAT	1322
Db	51539	ATACATATTTATATAGGCTTAACTAATCTGCTCCACCATATCTAGACTGTTTGTGAAT	514800
OY	1323	GAGTTGTCTATGCAATGTTTAAATTTTCAATTTTACATCTTTTCTGACAAAGCTTTTGA	1382
Db	51479	GGGTGTCTATACAGGTTTATATTTTGTAACTTTTATATCTTCTGAAAGCTTTTGA	514200
OY	1383	GTGGCCAAATATAGTCAGACATTCGAAAGTTTATGTGTATATATCATGATATTTTGAAT	1442
Db	51419	GTGGCCAAATATAGTCAGACATTCGAAAGGTTTATGTGTGTGTGTAGATATATTTGAAT	513600
OY	1443	GTAATAA 1448	
Db	51359	TTTAA 51354	

LOCUS	BD148636	807 bp	DNA	linear	PAT 17 -JAN-2003
DEFINITION	BD148636	Primer for synthesizing full-length cDNA and use thereof.			
ACCESSION	BD148636	GI:27854394			
VERSION	BD148636.1	GI:27854394			
KEYWORDS	JP 2002191363-A/3479.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Ota,T., Iwagaki,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002191363-A 3479 09-JUL-2002;				
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/3479 PD 09-JUL-2002 PE 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO, PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, PC C12P21/02, C12Q1/68 /C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key				
FEATURES	FT source 1..807 /Location/Qualifiers FT source 1..807 /Organism='Homo sapiens (human)'. Location/Qualifiers 1..807 /Organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'				
BASE COUNT	200 a 140 c 151 g 312 t 4 others				
ORIGIN					
Query Match	39.1%; Score 725.2; DB 6; Length 807;				
Best Local Similarity	95.4%; Pred. No. 1.6e-145;				
Matches	768; Conservative 0; Mismatches 34; Indels 3; Gaps 2;				
QY	444 CCTTATTATACAGACGAGCATCTATGTTTTCCTCTTTTGCCATATTTGATGCTCT				503
DB	1 CCTTATTATACAGACGAGCATCTATGTTTTCCTCTTTTGCGATTTGATGCTCT				60
QY	504 TTATGGAATCATTAACCTTCAGCCTTCCTGGAGTTTGGGCTTCATGTTGGCAAC				563
DB	61 TTATGGAATCATTAACCTTCAGCCTTCCTGGAGTTTGGGCTTCATGTTGGCAAC				120
QY	564 AAATATCATCTGGGCTGTCTTCTGTGCAGGGAGATGTCATTGCACAAAAGTTAACTGAGC				623
DB	121 AAATATCATCTGGGCTGTCTTCTGTGCAGGAAATGTCATTGCACAAAAGTTAACTGAGC				180
QY	624 TTGGAACACTAGCTACAAAGGAAGGAGACAGACTTCACCTATTAAGACCATTTGC				683
DB	181 TTGGAACACTAGCTACAAAGGAAGGAGACAGACTTCACCTATTAAGACCATTTGC				240
QY	684 AGAATTCGAAAAATCTTCAGTCTCTTTGGCTTATTCAGTGCCTTTAAACCTTGAG				743
DB	241 AGAATTCGAAAAATCTTCAGTCTCTTTGGCTTATTCAGTGCCTTTAAACCTTGAG				300
QY	744 TATGCTTTTCTGTGACTTGGCCCTACATCCTTCGGAGTTTCTGTTTGTGCTTTTGT				803
DB	301 TATGCTTTTCTGTGACTTGGCCCTACATCCTTCGGAGTTTCTGTTTGTGCTTTTGT				360
QY	804 AGTAGTTATGCTGGAATGTTTATTTGGGAGTCGAGTAGTCATGAAGCCTCTCTTCATT				863
DB	361 AGTAGTTATGCTGGAATGTTTATTTGGGAGTCGAGTAGTCATGAAGCCTCTCTTCATT				420

OY	864	TCCGACACTATTTCACCTTTTTCATTTATACCTCTTTTCTTCTTCGATCAGCTGC	923
	421	TCCGCAACTATTCTACCTTTTTTTTATTTACCTCTTTTTTTTCCCTTCATCTCTGC	480
OY	924	TCCTAGCAAAATTAAGACCTTTTCTTTCCTTAGTTGGAAACATGGAATTCGTTTTGGT	983
Db	481	TCTAGCAAAATTAAGACCTTTTCTTTCCTTAGTTGGAAACATGGAATTCGTTTTGGT	540
OY	984	GGTTACCTTAGTCTCTGTGTTTTAGTTTGGAAATTCATTATGCTCAATAATCTTGCCT	1043
Db	541	GGTTACCTTAGTCTCTGTGTTTTAGTTTGGAAATTCATTATGCTCAATAATCTTGCCT	600
OY	1044	AGCAGACAATATGACATTATTACTTTCCTATGTGTGAAAAAGATTTTTTCAAGATATGCAAT	1103
Db	601	AGCAGACAATATGACATTATTACTTTCCTATGTGTGAAAAAGATTTTTTCAAGATATGCAAC	660
OY	1104	TCTGAAATATTTTGTATGTTCCAGCCTATATATTTGCTGTGGTAGATAGCTGACTATT	1163
Db	661	TGT-AAATATTTGGTAGTTTCCAGCCTATATATTTGCTGTGGTAGATAGCTGACTATT	719
OY	1164	GAAATCAAAAGCCAAATTTTTTGGAAATTATAGTGTTCATATGCTT--GTTCATTGTATATA	1221
Db	720	GAAATCAAAAGCCAAATTTTTTGGAAATTATAGTGTTCATATAGCCTGGGTACAGTGATAG	779
OY	1222	GTTCCTCGAAGAACTGCTGGAAATTTCC	1246
Db	780	GTTCCTCGAAGAACTGCTGGAAATTTCC	804

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 11:15:44 ; Search time 44 seconds
(without alignments)
2533.171 Million cell updates/sec

Title: US-10-000-151b-3

Perfect score: 6079

Sequence: 1 MPVRGHAAPONTFLDTIIR.....GOLGALTSQPLRHGSDPGS 1159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6079	100.0	1159	2	I38465
2	1767	29.1	514	2	T18579
3	1614	26.6	1017	2	T13354
4	1531	25.5	1102	2	T17367
5	1537.5	25.3	1087	2	T13100
6	1485.5	24.4	1284	2	T13168
7	1448.5	23.8	1174	2	A40853
8	1409.5	23.2	962	2	I53197
9	1402	23.1	989	2	I46912
10	1401	23.0	934	2	T42394
11	451.5	7.4	7.4	2	A552072
12	440	7.2	7.2	2	A552072
13	433	7.1	695	2	S74179
14	428.5	7.0	688	2	S55349
15	425.5	7.0	735	2	I50630
16	423	7.0	732	2	S35691
17	412	6.8	787	2	S68699
18	409.5	6.7	645	2	I50680
19	409.5	6.7	663	2	S11521
20	409	6.7	686	1	A44842
21	407	6.7	682	1	JH0560
22	406	6.7	664	2	S11517
23	406	6.7	690	2	A42161
24	401	6.6	691	2	JC6509
25	398.5	6.5	688	2	B42161
26	397	6.5	688	2	S23606
27	397	6.5	857	2	S62694
28	393	6.5	883	2	T07651
29	391.5	6.4	677	2	S32816

30	391.5	6.4	690	1	S07103	CGMP-gated ion cha
31	385	6.3	887	2	T03939	potassium channel
32	384.5	6.3	697	2	H85205	potassium channel
33	384.5	6.3	697	2	T04931	potassium channel
34	382	6.3	828	2	I52046	potassium channel
35	379	6.2	575	2	I59327	olfactory cyclic n
36	375.5	6.2	772	2	S28292	hypothetical prote
37	375.5	6.2	880	2	F85381	potassium channel-
38	375.5	6.2	916	2	T05360	probable potassium
39	367.5	6.0	845	2	T07052	potassium channel
40	352.5	5.8	807	2	T12177	hypothetical prote
41	346	5.7	800	2	T19627	potassium channel
42	342.5	5.6	662	2	T04461	hypothetical prote
43	340.5	5.6	673	2	T20936	potassium channel
44	330	5.4	705	2	T06682	hypothetical prote
45	330	5.4	888	2	D84650	probable potassium

ALIGNMENTS

RESULT 1									
I38465									
probable potassium channel subunit - human									
C:Species: Homo sapiens (man)									
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 05-Nov-1999									
C:Accession: I38465									
R:Wanike, J.W.; Ganetzky, B.									
Proc. Natl. Acad. Sci. U.S.A. 91, 3438-3442, 1994									
A:Title: A family of potassium channel genes related to eag in Drosophila and mammals									
A:Reference number: A54953; MUID:94211879; PMID:8159766									
A:Accession: I38465									
A>Status: preliminary; translated from GB/EMBL/DBJ									
A:Molecule type: mRNA									
A:Residues: 1-1159 <RES>									
A:Cross-references: EMBL:004270; NID:9487737; PIDN:AAA62473.1; PID:9487738									
A:Superfamily: CAMP receptor protein cyclic nucleotide-binding domain homology									
F:742-858/Domain: CAMP receptor protein cyclic nucleotide-binding domain homology <CA									
Query Match									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MPVRGHAAPONTFLDTIIRKFE	GOSRKFIINARVENCAVYICNDGFC	ELGYSRAEYM	60				
DB	1	MPVRGHAAPONTFLDTIIRKFE	GOSRKFIINARVENCAVYICNDGFC	ELGYSRAEYM	60				
QY	61	QRPCTCDFLHGPRTORRAAQIAQAL	LGAEEKVEIAFYRKDGSCTCLVDVVPYKNE	DG	120				
DB	61	QRPCTCDFLHGPRTORRAAQIAQAL	LGAEEKVEIAFYRKDGSCTCLVDVVPYKNE	DG	120				
QY	121	AVIMEFLNFEVMEKDMVSGPAHDTN	HNGPPTSWLAPGRAKTFRKLPA	LALTARESSV	180				
DB	121	AVIMEFLNFEVMEKDMVSGPAHDTN	HNGPPTSWLAPGRAKTFRKLPA	LALTARESSV	180				
QY	181	RSGGAGGAGAPAAVVDVLTLPAP	SSSESLADEVTAMNHNHAGGPAER	RAIVPGSP	240				
DB	181	RSGGAGGAGAPAAVVDVLTLPAP	SSSESLADEVTAMNHNHAGGPAER	RAIVPGSP	240				
QY	241	PSASGQULPSRAHSLINPDASGSS	SLARTRSRESCASVRRASSAD	DIEMRAGVLP	300				
DB	241	PSASGQULPSRAHSLINPDASGSS	SLARTRSRESCASVRRASSAD	DIEMRAGVLP	300				
QY	301	RHAATGAMPRLSGLINSTDSDLV	RYRTISKIPOTTNLVFDLKGDP	ELASTPSDREITA	360				
DB	301	RHAATGAMPRLSGLINSTDSDLV	RYRTISKIPOTTNLVFDLKGDP	ELASTPSDREITA	360				
QY	361	PKIKERTHNVEKVOVLSIGADV	LPEYKLOAPRIHRMTILYSP	FKAWMDLITLIVY	420				
DB	361	PKIKERTHNVEKVOVLSIGADV	LPEYKLOAPRIHRMTILYSP	FKAWMDLITLIVY	420				
QY	421	TAVFTPYSAFLKETEEGPPATE	CGYACQAPVAVDLIVDIMFI	VDILINFTTVVNAE	480				
DB	421	TAVFTPYSAFLKETEEGPPATE	CGYACQAPVAVDLIVDIMFI	VDILINFTTVVNAE	480				


```
Db 133 ISOSGGPLGSGS---IHGD-----NNNHENSLG-----RR-----160
QY 240 PRSAPGQLPSPRAHSLNPDASSGSLARTRESCASVRRASADIEAMAGVLP 299
Db 161 -----GASSRLNSTRQN-----173
QY 300 PRHASTGAMHPLKSLNSTSDSLVRYTRISKIPQITLNFVDLKGDFLAPTSREII 359
Db 174 -----RTV--LHRLTGHF-----GRDQGSVKANSVF 199
QY 360 APRKIKERTHNVTEKVTQVLSLADVLPYKQLAPRIHRWTLHYSPFKAWMDLILLYI 419
Db 200 EPK-----PSVPEYKVASVSGSCRLILLYISIPKAWMDLILLYATF 239
QY 420 YTAFTPTPSAFLKETEGBGPATCEGYACQPLAVVDLYVDIMFYDILINFTTYVNNAN 479
Db 240 YVAVTYPIYVNCF---AGDDTPITS-----RHTLVSDIAEMLEFIDILINFTTYVSOS 291
QY 480 EEVYVSHRGRIVAHYFKGWFLLDMVAIIPDIL-IFSGSSEELIGLKTARLLRVAR 538
Db 292 GGVVSAIPRSIGLHYLATWFEVDILALPDLLVFNITVSLVHLKTVRLRLRLK 351
QY 539 LDRYSEGAAYFLMCTPALLAHMLACTIWAIGNMDPHMS---RIGLHNLGDOIG 595
Db 352 LERYSQCSAVVLTLLMSVALLAHMMAWYVIGRREMEANDPLMDIGLHLEGRLEE 411
QY 596 PYSSGLGSPSIXDKVYVLTALYFESSLTSGVGNVSPNTSEKIFISCVMLIGSLMVASI 655
Db 412 PYVNGSAGGPRSRSAVIAALYFTLSSLTSGVGNVSPNTSEKIFISCVMLIGSLMVASI 471
QY 656 FGNVSAIIORLYSGTARHTOMLRVREFIRHOIIPPLRQRLSEYFOHAMSNTNGIDMA 715
Db 472 FGNVTAIIORMYSRSLYHSRMDKDFIRVHRLPRLKQRMLEYQOTWAVNSGIDANE 531
QY 716 VAKGFEPCLOADICHLNLSLOHCKPFGATKGCIRALAMKKTTHAPDGLVHAGDL 775
Db 532 LKDFPEDELRAIDAMHNLREILQ-LPLFGASRCGLRALSHIKTSCAFCEFLLRGDA 590
QY 776 LPAIYISRGSIETILGADVVALTGKNDIFGEPL-NLYARPG-----KSNDDVALTY 827
Db 591 LQAHYVYSGSLSELDNDVTLALGDLIGADIPELGQEPGAGACVLTSDVVALTY 650
QY 828 CDIAKIHRLDLEVLDMYEFSDHFWSSL-EITFNLRTNMIPGSPGSTELEGFSR-Q 884
Db 651 CGIQQLSSRGLAEVLLPEYVAARAGLPROLITFLRQ-----GSENN-----GLGRS 700
QY 885 RKRKLSIFRRR---TDKDTQEGEVALGPRAGAPSSRG-----RPGGFWG-----928
Db 701 RSPRLSQARSDDLGSSSDKTLPISITETEGMEBAGSKPRRPLLPNLSPARRGSLVSL 760
QY 929 ---ESP--SSGSPSPSSSEDECP---GRSSP-LRLVFPSSP---RRP-----GEP 967
Db 761 LGEELPFPFALVSPSPSPSPALAGRSSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 820
QY 968 PGGEPIMEDCEKSSDPCNPGLSCAFSGVSNIFSEMGDSRGROYELRCPAPPSLNLPI 1027
Db 821 DLSPMTVVDGEDSSNTAEAPTRFS-----KREPTITRQAPL 859
QY 1028 SSP-----GRPRGVESRLDALQOLNLE---TRLSDAMATVLOLQORMTLVPAY 1078
Db 860 SGRLRELATEAEAEVEKVKVCRLNQELISRLNQEVSQLSRELQVGLQAR-LGPPSH 917
QY 1079 SATYT-----PGGPTSTSLPLVPSPPLTILDSLSQVSOQMAC 1117
Db 918 PPDSTYLPDLPCPHQRPCTISPMHSGPPPLQMTTLLAVVHCPASVGTVEIGATPSE-LRS 976
QY 1118 EELIPGAPELPQEGPTRRLSLPGOLGALTSQLPHRGSDDGS 1159
Db 977 SMVPPPPSPEDPLGPS-----PYPEASPLPLSLIKHSFQSGS 1013
```

RESULT 4
T17367

```
potassium channel protein elki - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17367
R:Shi, W.; Wang, H.S.; Pan, Z.; Wymore, R.; Cohen, I.S.; McKinnon, D.; Dixon, J.E.
J. Physiol. 511, 675-682, 1998
A:Title: Cloning of a mammalian elk potassium channel gene and EAG mRNA distribution
A:Reference number: Z18731; MUID:98382545; PMID:9714851
A:Accession: T17367
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1102 <SH>
A:Cross-references: EMBL:AF061957; NID:93659689; PID:93659690; PIDN:AAC61520.1
C:Genetics:
A:Gene: elki
C:Function:
A:Description: may play a role in the sympathetic nervous system
C:Keywords: potassium channel

Query Match 25.5%; Score 1551; DB 2; Length 1102;
Best Local Similarity 32.6%; Pred. No. 7.9e-90;
Matches 383; Conservative 169; Mismatches 283; Indels 340; Gaps 30;

QY 1 MPVRGHVAPQNTFLDIIIRKFGQSRKFLIANARY-ENCAYITCNDGFCGLGYSRAEV 59
Db 1 MPVKGGLAPQNTFLDIIATRFDGTHSNFLANAQYAKGPIYYCSDGFCGLGAFRTYV 60
QY 60 MQRPTCDPLHGRTRORRAAQAQALLCAEEKVELATYRKQSGFCLLVDPYVKNED 119
Db 61 MOKSCCKPLFGETNEOLMLQTEKSLSEKVEKPEKEMFKKNGAPFCMLDITVPIKNEK 120
QY 120 GAVIMTFLNEVMEKDMGSPAHDPNNHRRPPTSWLAPGAKTFRLKLPALLALTARESS 179
Db 121 GDVYLFASF---KDIITDKY-----138
QY 180 VRSGAGAGAPGAVVVDVLTLPAPSSSLADEVTAMDNHVAIGLPAEERRALVPGS 239
Db 139 -----KITSDEK-----KEDR-----149
QY 240 PRSAPGQLPSPRAHSLNPDASSGSLARTRESCASVRRASADIEAMAGVLP 299
Db 150 ---AKGR-----SRASHPDSARRSR-----168
QY 300 PRHASTGAMHPLKSLNSTSDSLVRYTRISKIPQITLNFVDLKGDFLAPTSREII 359
Db 169 -----AVLYHISGLQREKKKLTINNVEFDK-----197
QY 360 APRKIKERTHNVTEKVTQVLSLADVLPYKQLAPRIHRWTLHYSPFKAWMDLILLYI 419
Db 198 -----AFPEYVSDAKSKSKFTLHFSFKAGMDLILLYATF 233
QY 420 YTAFTPTPSAFLKETEGBGPATCEGYACQPLAVVDLYVDIMFYDILINFTTYVNNAN 479
Db 234 YVAVTYPIYVNCFIGNEDLSTTRST-----TVSDIAEILFIIDILINFTTYVS 284
QY 480 EEVYVSHRGRIVAHYFKGWFLLDMVAIIPDIL-IFSGSSEELIGLKTARLLRVAR 538
Db 285 GOVIFEARSTCIHYVTTFITDILALPELALFAENVTVVSLVHLKTVRLRLRLK 344
QY 539 LDRYSEGAAYFLMCTPALLAHMLACTIWAIGNMDPH---MDSRIGLHNLGDOIG 595
Db 345 LDRYSQHSSTVTLNLSMFLAHMMACTIYVGNKKEREDNSLKMEVGLHLEGRLES 404
QY 596 P-YNSGGLGSPSIXDKVYVLTALYFESSLTSGVGNVSPNTSEKIFISCVMLIGSLMVAS 654
Db 405 PYVGNNTLGGPSIRSAYIALYFTLSSLTSGVGNVSPNTSEKIFISCVMLIGALMAL 464
QY 655 IEGVNSAIIORLYSGTARHTOMLRVREFIRHOIIPPLRQRLSEYFOHAMSNTNGIDM 714
Db 465 VFGNVTAIIORMYSRSLYHSRMDKDFIRVHRLPRLKQRMLEYQOTWAVNSGIDSN 524
QY 715 AVLKGFPECLQADICHLNLSLOHCKPFGATKGCIRALAMKKTTHAPDGLVHAGD 774
Db 715 AVLKGFPECLQADICHLNLSLOHCKPFGATKGCIRALAMKKTTHAPDGLVHAGD 774
```

Dh 525 ELKDPDELRSDDTMMHINKNEIIQ-LSLFECASRGCLRSLSLAIKTSFCAGEXEYLLROGD 583
Qy 775 LITAFYISKGSIEILRGDVVAIIAGKNDJFGEPLNYARPGKNGDVRATYCDLKH 834
Dh 584 ALQATYVSGSMYKLVKDSMTALIGKDLIGANLSIKDYIKTNADYKALTCDLQCTI 643
Qy 835 RDLLLEVLDMYPEFSDFHFWSSL-EITFNILND-----TNMIPGSGSTEELEGG 880
Dh 644 LKGLFEVIGLYPEYAKHFVEDIQHDLTYNLREGHESDVISRLSKSTYV-----QAEPRGN 699
Qy 881 FSRKRKKLSFRRTDVTDTQDGEVSALGPRACAGSSRGRPGMGESPPSGSPSPES 940
Dh 700 GS-LKKRLPSIVDEDEEEVEEETSLSP-----LYTRG-----SSVSH 738
Qy 941 SEDGPRSSSPLRL-----VFPSSP-----RPPGEP----- 968
Dh 739 SKTKGSSKSYIGLSLKQLTSGTVFHSPIRYSANSKRTKQEOADPPNHGTRKEKNLVOL 798
Qy 969 -----GGEPLM-----EDCEKSSDTCNPLSGAFSGVSNIFSPWDSRGRQYQELPRC 1015
Dh 799 CSLGTAGTPELSPIRYDGIEDGNSSEET-----QTFDFGSE---QIRPEPRI 842
Qy 1016 PAPPSLNLPLSSPGRRPRQDV-----SRDALQRLNLE-----TRISA 1058
Dh 843 ---SPSL-----GESEIGAFLFIKAETKQOINKLNSEVTLTLOESOLGK 886
Qy 1059 DMATVQLQLOROMTLVPPAYSAVTPPGPPTSTSP 1093
Dh 887 DMRSIMQLENIILSPQPSORC-----SLHPTSTCP 917

RESULT 5

311100
probable potassium channel 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: J11100
R:Engelard, B.; Neu, A.; Ludwig, J.; Roepert, J.; Pongs, O.
Submitted to the EMBL Data Library, July 1998
A:Description: Identification of three rat potassium channel genes homologous to D. melanogaster
A:Reference number: 220983
A:Accession: J11100
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1087 <ENG>
A:Cross-references: EMBL:AJ007627; NID:e1329995; PID:e1329996; PIDN:CA07586.1
A:Experimental source: cortex
C:Genetics:
A:Gene: elk2
C:Keywords: potassium channel

Query Match 25.3%; Score 1537.5; DB 2; Length 1087;
Best Local Similarity 32.4%; Pred. No. 5.5e-89; Indels 373; Gaps 35;
Matches 419; Conservative 143; Mismatches 357;

Qy 1 MPVRGHVAPONTFLDTIRKFEQSRKFIANARVENC-AVIYCNDFCELCYSRAEV 59
Dh 1 MPARGLLAPQNTFLDTIRKFEQSRKFIANARVENC-AVIYCNDFCELCYSRAEV 60
Qy 60 MQRCTCDLHGPTORRAAQIAQALLGAERKVEVLAFTYRKDSCFLCLVDVVPKNE 119
Dh 61 MQRCTCDLHGPTORRAAQIAQALLGAERKVEVLAFTYRKDSCFLCLVDVVPKNE 120
Qy 120 GAVIMFLINFEVMEKMWGSPADTNHRGPTSMVLAAGRAKTRKLRKLPALALTARSS 179
Dh 121 GEVLEFL-----VSHKDI-----SETKNRGGRPDWMEKGGRR----- 153
Qy 180 VRSAGAGAGARGAUVVDVLTTPAARSESLALDEVTAMDHNVAAGLGAEEERRALVGP 239
Dh 154 -RYGRASGK----- 162
Qy 240 PPRSAPOQLDSBRAHSLNPDASGSSCLARTRSRESCASVRRASADDIEMRAGVLP 299

Dh 163 -----FNANRRRSFVLYHL-----SGHLOQ 184
Qy 300 PRHASTGAMPLRLRGLNSTSDSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREI 359
Dh 185 PK-----GKHKLNGVGEKPN----- 201
Qy 360 APKIKERTHNVEKVTQVLSIGADVDEPKYLOAPRIHRMTILAHSPKAWMDLILLLVI 419
Dh 202 -----LPEYKAAIRKSPFILLHCGALRAWDGFIILATL 236
Qy 420 YTAFTPYSAFLKTEDEGPATEGVAQOPLVAVDLYDIMFTVILINFRTTYNAN 479
Dh 237 YVAVTVPSYCV-----STAREPSA-----AKPPSCDLAAVEFLIDIVLNFRTFVSKS 288
Qy 480 EEVVSHGRVLAHYHFKMFLIDVVAIIPDIL-IFGSGSEELIGLITARTLRVYARK 538
Dh 289 GQVVFARSKISCIHVTWFLPDVIAALPFDLHNFKNVYVGAHLITVRLRLRLR 348
Qy 539 LDRYSEGAVALFLMCTFALIAHMLACIWAYIAGNMEQPHMDS---RIGWLHNGDQIG 595
Dh 349 LDRYSQSAVYVLTLLMAVFALLAHVAVCFEYIQOGEIENSESELPEIGWLOELARLET 408
Qy 596 PY-----NSSG-----LGGSTIKDKYVTLALYTFSSLT 624
Dh 409 PYYLVSRSPDGNSSGSENCSSGSGSEANGTLELGGPSLSAVITSLYFALSST 468
Qy 625 VGFGNVSPNTNSEKIFICVALLISLWYASIFGNVSAIIQRYLSTGRYHTOMLRVEFI 684
Dh 469 VGFGNVSNANDTEKIFISCTMIGALMHAVYGVNTAIQRMVARRLYNSRTDLADI 528
Qy 685 RFHQINPDLROLEVEYOHAMSYTNGIDMAVLEKGPPECIOLADICHLNLSLOHCKPFR 744
Dh 529 RIRHPRKPKRMLEYEQATWAVNNGIDTTEILQSLDELADIAHMLHKEVLO-LPLFE 587
Qy 745 GATKGCCLRALAMKKTTHARPQDVLHAGDILTALYFISGSEILIGDVAVALIGKNDI 804
Dh 588 AASRGCLRALSLARPACTPGEVLIHQGDALQVLYVSCSMVLELGGVLYLIGLGD 647
Qy 805 FGEPLNLYARPGKNGDVRATYCDLKHIRHDLLEVLDMYPEFSDFHFWSSL-EITFNIL 862
Dh 648 IGCGLPQREQVKNNAVDKGLTYVLCLOLAGHESIALYPERVAPRFSRGLRELSTY 707
Qy 863 R-----DTNMIHG-----SPGSTEELEGG 880
Dh 708 GAGGVSAEVDTSLSGDNLTLMSTLEKETDGEQHTISPADEPSSPLSPGCT----- 762
Qy 881 FSRQRRKLSFRRTDVTDTQDGEVSALGPRAG-----AGPSSRGR-----PGGPMG 928
Dh 763 -SSSSAAKLSPRT---APRPLRGGRGRSRAGVLRKPGPSAHPRTLDGLQLPRPMWN 818
Qy 929 ESPSGSPSPESSED-----EGPGRSSSPLRLVFPSSPRPGEPPG---GE 971
Dh 819 VPPDLSPRYVDGIEDGGSDQHKRSPRYGGSGPCCSSP-----SPGTESGILTV 868
Qy 972 PLMEDCKSSDTCNPLSGAFSGVSNIFSPWDSRGRQYQELPRCAPPSLNLPLSSPG 1021
Dh 869 PLVPSSEARNEDTLDKLOAVTELS-----EVOYLMREIDQSLRAVOVLIVP-QGEG 919
Qy 1032 RRRP-----GDVESRLDLQRLNLETRLASDMATVLIQILQRLQMTLVPPAYSAV--TTPG 1085
Dh 920 QCPRVSGEGPCPAVASGLDPL-RVDYGAAS-----YCLDPPASVLSGTWPH 966
Qy 1086 PGPTSTPPL-----PVSPLP-----TLLDLSLQVSOFMACEELPRCAPELPQ 1129
Dh 967 PRGCHPRPLMAPWGWGPRASGSSPWRATYALMTITSDEPPSGSDLCSESPSTASPPPE 1026
Qy 1130 EGPTRLSLPQGLALTS---QPLRHGSDPGS 1159
Dh 1027 EGARTGTAPAVSQAEATSTGEP-----PPGS 1052

RESULT 6

T13168

probable potassium channel elk chain - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13168
R:Wanke, J.W.; Ganetzky, B.
Proc. Natl. Acad. Sci. U.S.A. 91, 3438-3442, 1994
A:Title: A family of potassium channel genes related to eag in Drosophila and mammals.
A:Reference number: A54953; M0ID:94211879; PMID:8159766
A:Accession: T13168
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1284 <MAX>
A:Cross-references: EMBL:U04246; NID:g487735; PID:g487736; PIDN:AAA62472.1
C:Genetics:
A:Gene: elk
A:Cross-references: FlyBase:FBgn0011589
A:Map position: 2R

Query Match 24.4%; Score 1485.5; DB 2; Length 1284;
Best Local Similarity 32.3%; Pred. No. 1.3e-85;
Matches 366; Conservative 173; Mismatches 330; Indels 265; Gaps 27;

QY 1 MMYRGRHVAPOFTLDTIRKEFGGSRKTIANARVENCAYICNDGFCGLGYSRAEYV 60
DB 1 MPARKGLAPONTFLDTIRKEFGGSRKTIANARVENCAYICNDGFCGLGYSRAEYV 60
QY 61 ORPCTDPLHGPRTORRAAQAIALGAERKVEIAFYRKOSGFCGLVDVYVKNEDG 120
DB 61 QKGCCHFLYEPRTKEHKOJLEKSNMELKLEYIFKKEBAPWCPLDYPITNEKR 120
QY 121 AVIMFLNEVYMEKDVSPAHDTNHRPPTSMALPGRAKTFRLPALALTAARESSV 180
DB 121 DVLFSLASHKIDITHTKMLEMNVNE-----ECDSVFALTAALGA 159
QY 181 RSGGAGACAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAAGLCPAERRALVPGSP 240
DB 160 RFRAGSNAG-----MLGIG-----GLPGLGSP 181
QY 241 PRSAPQLPSRAHSLNPDASGSSCLARTSRSCASVRRASADDEAMRAGVPPPP 300
DB 182 AAS-DQDTAGCGNNIDVPA--GCMGRRSR-----AVL----- 213
QY 301 RHASTGAMHPLNSGLINSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITA 360
DB 214 -YQLSGHYKPEKGV-----KTKLKLGN--NFMHSTEAPF----- 245
QY 361 PKIKERTHNVTEKYOVLSGADVLPYKIQAPRIHRTITLHSPKAVMDLILLYIY 420
DB 246 -----PEYKTOSIKSRLLPHYGVKGIWDVILVATFY 280
QY 421 TAVFPPYSAFLKTEEGRPATEGACOPLAVVDIYDIMEITVILNFTTYVNAE 480
DB 281 VALMVPYNAF-----AKADQIKVSIVYIEALFIVILNFTTFEYSKRG 326
QY 481 EYVSHGRIAVHYFKGMFLIDVVAALIPFDLLIFG--SGSEELIGLLKTAARLRLVAVR 537
DB 327 EYVSNKOJAALNYLNGFALDLAALPFDHLVASDLYDGSDSHILVKTLRLRLARLDQ 386
QY 538 KIDRYSEYGAALVFLMCTFALIAHLACIWAIGMEOPHM-DSKIGMLHNLGDDIGRP 596
DB 387 KIDRYSOHTAMITLTLMFSEFTLAHMLACIWIYIAKEYEWPESNIGMQLAER---- 442
QY 597 YNNSGIGGSPSIKDKYATYALTFEFSSTSVGFNVSPTNSEKJFSCVMILGSLMASTF 656
DB 443 -KNASVAILTAEYISTALYFTFTSLTSGVFGVNSANTAEKYFTIIMLIGLMAVAF 501
QY 657 GNVSAITQRLYSGTARYHTQMLKVRREFIRPHQIPNDLQRLLEEYFQHANSTYNGIDMANV 716
DB 502 GNTVATIQRMYSRSLYESKMRDLKDFVALHNNPKELKORIEDYFQTSMSLSHGIDYET 561
QY 717 LKGFPPCLQADICLHNBSLQCHKRFPRATKGCCLALAMKFTTTPAAPGCDTLVHAGDLL 776
DB 562 LRFPEELKGDVSMHLAREILO-LPIFEAASQCLTLSLHKTNCPACAGEYLIRHGDA 620

QY 777 TALYFISRSIELRGDVVAALIGKNDIFGEPLNLYARPKSNG----- 820
DB 621 NYIYLCNMSMEVETKIDMVAALIGKNDIVGSDINVL-VATNSGQMTATNTNSAGDVYVR 679
QY 821 ---DVALRYCDLHKITHRDLEVDLMDYEPFSDHFWSSL--ETTFNLR-----DTNMT 869
DB 680 SSSDIALYICDIDKCIHMGGLVEVLRLYPEYQOOFANDIONHDTICMLREYENODSDIG 739
QY 870 GSPGSELEGGFGRQKRKLSPFRRTDKTEQGEVSALGPAGAGSPSSRGPGPWE 929
DB 740 SEPLPPISE-----DDEKREAEKGGKEKENGSGPBGASPLHNISN 782
QY 930 SPSSGSPSSSEDEGPRSSSPRLVFPSSPPRPEPGGEPL--MEDCEKSSDYCNPL 987
DB 783 SPLHATRSF-----LLGMSPRNQLHQRGRSLITLRETNKHNRLNMA 826
QY 988 SGAFSGVSNIFSWGDSRGROYELRCPAPPSLNLPLSSPGRRRPDGVSRLDALOR 1047
DB 827 CSLDRG-----SFE-----PEPLEEQQSGCKRPSLE--RLDSQYS 861
QY 1048 QLNRLTRLSADMATVQLQRLQMTLV-----PRASVATTTGPGPT 1089
DB 862 TLHQDVAAQLSAEVRNALISALQ--EMTETSNAMTSHSSIKRPPARSIPNISGAGT 914

RESULT 7
A40853
potassium channel protein eag - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 24-Sep-1998
C:Accession: A40853
R:Wanke, J.; Drysdale, R.; Ganetzky, B.
Science 252, 1560-1562, 1991
A:Title: A distinct potassium channel polypeptide encoded by the Drosophila eag locus
A:Reference number: A40853; M0ID:91262635; PMID:1940699
A:Accession: A40853
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1174 <MAX>
A:Cross-references: GB:M61157; NID:g157311; PID:g157312
C:Genetics:
A:Gene: FlyBase:eag
A:Cross-references: FlyBase:FBgn0000535
C:Keywords: transmembrane protein

Query Match 23.8%; Score 1448.5; DB 2; Length 1174;
Best Local Similarity 30.4%; Pred. No. 2.6e-83;
Matches 374; Conservative 174; Mismatches 328; Indels 355; Gaps 33;

QY 4 RGHVAPONTFLDTIRKEFGQ-SRKEITANARVENCAYICNDGFCGLGYSRAEYV 62
DB 5 RGLVAPONTFLDNITRRSNSQDSSFTLANAOIYDPPITYCESFCISKINRAEYVOK 64
QY 63 PC--TCDPLHGPRTORRAAQAIALGAERKVEIAFYRKDG--SCFL----- 107
DB 65 SCRYVCGFMKGELTDEKREYGLREYLENQOQDFEILYKKNMLQCGCALSGRGAQTOE 124
QY 108 ---CLVDVVPKKNEGCAVIMFLNEVYMEKDVSPAHDTNHRPPTSMALPGRAKTF 163
DB 125 TPLMLLOVAPYIENEDLVVFL-----TF 150
QY 164 RLKLPALALTAARESSVRSAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVA 223
DB 151 R-----DITALKOPT- 160
QY 224 GLGPAERRALVPGSPRSAPQLPSRAHSLNPDASGSSCLAR--TRSRSCASVRR 281
DB 161 ---DSBDTGVLG-----LSKFAKLARSYTRSRQFSAML-- 191
QY 282 ASSADDIEAMRAGVLPFRHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNV 341
DB 192 ----- 191

[illegible]

RESULT 8
153197
potassium channel subunit - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C:Accession: 153197
R:Luftwig, J.; Terlau, H.; Wunder, F.; Bruggemann, A.; Pardo, L.A.; Marguardt, A.; Stuhme
EMBO J. 13, 4451-4458, 1994
A:Title: Functional expression of a rat homologue of the voltage gated ether a go-go po
ts mammalian counterpart.
;Reference number: 153197; MUID:95009946; PMID:7925287

A:Accession:U53197
A:Status:Preliminary; translated from GR/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-962 <RES>
A:Cross-references: EMBL:U34264; NID:9557264; PIDD:CA84018.1; PID:9557265

Query Match 23.2%; Score 1409.5; DB 2; Length 962;
Best Local Similarity 30.9%; Pred. No. 5,76-81;
Matches 338; Conservative 151; Indels 329; Gaps 23;

[illegible]


```

QY 902 -----PGEVSALGPGRAGCPSSR----- 920
Db 759 SVTVRESPATPVSFQAASSTVSDHAKLHAPGS - ECLGPKAGGADPAKKGNAFRKDAC 817
QY 921 GRGGGWGE-----SPSSGPSSPESEDEBGPGRSSSPLRLVPFSSPRPPGEP 967
Db 818 GK-GEOMNKVSKAESMETLPERTKASGEATLKTDCSDSGITKSDRLDNVNGEARSPD- 875
QY 968 PGEPLMEDCKS 980
Db 876 --RSPILAEVYKHS 886

RESULT 9
148912
potassium channel subunit - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: 148912
R:Warme, J.W.; Ganetzky, B.
Proc. Natl. Acad. Sci. U.S.A. 91, 3438-3442, 1994
A:Title: A family of potassium channel genes related to eag in Drosophila and mammals.
A:Reference number: A34953; MUID:94211879; PMID:8159766
A:Accession: 148912
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-989 <RES>
A:Cross-references: EMBL:U04294; NID:9487739; PIDN:AAA62474.1; PID:9487740
C:Genetics:
A:Gene: m-eag

Query Match 23.1%; Score 1402; DB 2; Length 989;
Best Local Similarity 29.7%; Pred. No. 1.8e-80;
Matches 361; Conservative 166; Mismatches 363; Indels 322; Gaps 27;

QY 4 RGHVAPONTFLDTIRKFEQGRKFIANARVENCAYIYNDGFCELGYSRAEVQRP 63
Db 7 RRLVLAQONTFLFNVIYR--SNDTNFVLGNAQIYDMPYIVSNDGFCLSGYHRAEVQKS 64
QY 64 CTCDLHGPRTORRAAQIQAALLGAERKEVETAFYKDKDSCFLCYLDVYPRVKNEDGAVI 123
Db 65 SASCFMVGELTDQYVEKVRQTFENYEMNSFEILMKKRNTPYWFYKIAPIRNEODKV 124
QY 124 MFLINEFVMEKDMVGSPPADTNHRGPPSTMLAPGRAKTRFLKLPALLALTARESSVRS 183
Db 125 LFLCTRS----- 131
QY 184 GAGGAGAGAVVVDLTTPAAPSSSLALDEVTAMDNHVALGSP-AEERRALVGPSPPR 242
Db 132 -----DITAFKQPLE-----DDSCKGWGFALUTRAL----- 158
QY 243 SAGGQLPSPRAHSLNPDASSSSCSLARTSRSCASVARRASADDEIAMAQVLPPEPRH 302
Db 159 -----TSSR----- 165
QY 303 ASTGAMHPLRSGLNSTSDSLVRYRTISKIPQTLTNFVLDKGPFLASPTSDREITAPK 362
Db 166 -----QQLAPS 171
QY 363 IK--ERTHNTEKVTQVLSGADVLPPEYKLAQRIHRTILHSPKAVWDMLLLVIY 420
Db 172 VQKGENVHK-HSRLEAVLIQSGDILPQYKQEARPTPHILIKCVFTTMDWILLITFY 230
QY 421 TAFPTPSAFLAKEKEEGPRATECGACQPLAVVDLIYIMFVLDLIMFRTTYVANE 480
Db 231 TALLVPNVSEFKRONNA-----WLVVDSIVDVIVFLVDIVLNFHTTFVGPAG 278
QY 481 EVVSHPRIVAHYFKGFLIDVVAIPEDL-----IFG----- 514
Db 279 EVISDRKLIMNLTKTFTVIDLISCLPYDIINAFENVDEVSAFGDPDGKIGFADQIPPL 338
QY 515 -----SGSEBLIGLKTARLLRLVVARAKLDTRYSEGAVALFLMCTFALIAHMLACIMY 569

```

```

Db 339 EGRESQISSLSFSLKLVRLRLGRVARKLDHYIEYCAAVLYLVYFGLAHMACIMY 398
QY 570 AIGMNEQPHMDSRI-----GWLHNLDQIGRPY--NSSG-----GGPSIKQYVALKFTF 619
Db 399 SIGDYELFEDDTYKTRINNSWLYOLALDIPPOYONGSGSKWEGPSKNSVYISLSYFTM 458
QY 620 SLSLVGFGVNSPPTNSEKIFSLICVMLTGLSMVASTIGVNSAITORTSGTARYHTOMLR 679
Db 459 TSLTSGFGNIAPSTOLEKIFPAVAINIGSLTYITFGNTTIFQOMYATNKRHEILNS 518
QY 680 VREFIRPHQIPNPLRQLEEYFOHAWSYTNGIDMNAVLKGFPECLQADICLILNRSLLQ 739
Db 519 VRDFELKLYQVPKGLSERVMYIVSTWMSMGIDTEKVLQICPKMDRADICVHLNRKVEKE 578
QY 740 CKPFRGATKGLRALAKFKTTIARPDQTLVHAGDLTALYFSRGSIELRGVVAAL 799
Db 579 HPAFRLASDCLRLALAMEFQTVHCAPDQLTYHAGEVSYSLCFVWSGLEIYODEVVAAL 638
QY 800 GKNDIFGEPLNLYARPGSNGDVRLATYCDLHKITHRDLLLEVDMYEPESDHEWSLEIT 859
Db 639 GKGDVEGDVWKEKTIILAQSCANRALTYCDLHVKRALQKVEFTYAFSHSPSRNLT 698
QY 860 FNLR-----DTN-----MTPSGPSTLEEGFSRQRRKLSFR-- 893
Db 699 YNLKRLIVFKRISDVKREEREMKRNKEAPLILPHDPVRRLFORFROKEARLAERGG 758
QY 894 --RTDKTDEQGEVSALGPGRAGPSS-----RRPGCPMGESSPSSPESEDEBGP 947
Db 759 RDLDDLDVEKGNALTDHTSANHSLVKASVYTVRSPTPVPSFOAATYSTSDHAKLHAPG 818
QY 948 RSSSPLRLVPFSSPRPPEPGEPLMEDEKSSDTGNPLSGAFSGVSNIFSEWGD-SRG 1006
Db 819 SEC-----LGRKAVASCDPAKKGNAFRKDACGKD-----WNKYSKA 856
QY 1007 ROYQELP-RCPAPPSLLNIPLSSGRRPGDVESRLDALQRLNRLTSLADMAVQLQ 1065
Db 857 ESMETLPERTKARPEATLKTDCSDSGITKSDL--RLD-----NVGETRSPQORSPILA 908
QY 1066 LLDROMTLVPRAVSATVTPPGCPRTSPPLVPSPLPRLTLDLSOVQOPMACELLPFGAP 1125
Db 909 EVKHSFYPIPEQTLQAAVLVLYKELKEDIKALNAKMYSIEKQISEILLRIL---MSRGA 964
QY 1126 ELPOEGPTRLSLP 1139
Db 965 QSPQE--TGETISR 976

RESULT 10
142394
potassium channel protein eag homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T42394
R:Waterston, R.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z22153
A:Accession: T42394
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-934 <NMT>
A:Cross-references: EMBL:AF036695; PIDN:AA88348.1
C:Genetics:
A:Introns: 27/1; 66/2; 104/1; 147/1; 164/3; 222/2; 263/3; 301/3; 410/1; 468/2; 667/3;
A:Note: F10B3.1

Query Match 23.0%; Score 1401; DB 2; Length 934;
Best Local Similarity 29.6%; Pred. No. 1.9e-80;
Matches 352; Conservative 167; Mismatches 304; Indels 366; Gaps 27;

QY 1 MPV-RGHVAPONTFLDTIRKFEQGRKFIANARVENCAYIYNDGFCELGYSRAEV 59
Db 1 MPVGKRLVAPONTFLFNVIYRRCNNADTSEFLANAQVVDYIVYCNDFSGKIVGYTRAEL 60

```


C:Species: Bos primigenius taurus (cattle)
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 13-Aug-1999
 C:Accession: A55251; S43976
 R:Biell, M.; Zong, X.; Distler, M.; Bosse, E.; Klugbauer, N.; Murakami, M.; Flockerzi, V.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3505-3509, 1994
 A:Title: Another member of the cyclic nucleotide-gated channel family, expressed in testis
 A:Reference number: A55251; MUID:94224768; PMID:8170936
 A:Accession: A55251
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-706 <BIE>
 A:Cross-references: GB:X76485; NID:9488728; PIDN:CAA54023.1; PID:9488729
 A:Experimental source: kidney
 R:Weyand, I.; Godde, M.; Frings, S.; Welner, J.; Mueller, F.; Altenhofen, W.; Hart, H.;
 Nature 368, 859-863, 1994
 A:Title: Cloning and functional expression of a cyclic-nucleotide-gated channel from man
 A:Reference number: S43976; MUID:94211295; PMID:7512693
 A:Accession: S43976
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-706 <MEY>
 A:Cross-references: GB:X89600; NID:9908023; PIDN:CAA61759.1; PID:9908024
 A:Experimental source: testis
 C:Genetics:
 A:Gene: CNGB3
 C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-
 C:Keywords: cAMP binding; ion channel; ion transport; membrane protein
 F:501-625/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>
 Query Match 7.28; Score 440; DB 2; Length 706;
 Best Local Similarity 23.3%; Pred. No. 6.3e-20;
 Matches 162; Conservative 115; Mismatches 283; Indels 134; Gaps 21;
 QY 191 PGAVVVDL-----TPAASSESLADEVTADNVAAGPAREERALLVGPSP 241
 DB 16 PSVTRMDRDLDCTENGLSRHLPCETSSSELQEIAMETR---GLAESRQSSFTSGPT 71
 QY 242 RSAPGOLSPRASH--LNPDASSGSCSLARTSRSCASVRRASSADDEAMRAGVLP 298
 DB 72 RLISRLASARLHEDQRPDSFLERFG--AELEQSSSHSHVQFVGSQEP 126
 QY 299 PRPHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPPLASPTDREI 358
 DB 127 PDGRSAMPALARNTNTNNSSEKDDKAKKEKEKE-----EKKENPKKEKKDSV 179
 QY 359 IAPRIKRTNHTVEKYQVLSIGADVLPEYKLAQRIHRWTLIHSFFKVM----- 411
 DB 180 MDSSSNMYHWLT-----VIAVPFTNMCL--VCRACFDELQSEHL 219
 QY 412 --WILLLVITYAVETPYSAFLKETEGRPATECGYACOPLAVVDLIVIMEYVILI 469
 DB 220 MLMLVL-----DYSA-----DILGMDMLV 239
 QY 470 NFRTTYNANEVYSHGRIVAFYFG-WFLIDMVAIIPDDLIFSGSEELIGLKTAR 528
 DB 240 RARFGLEQG-LVMWASRLMKHYQTLHFKLDVLSVPPDLAVFKGMN--YPELFPNR 296
 QY 529 LRLVRARKLDRS-----EYGAVLELLMCTFALIAHMLACIWAIGMEQPHM 579
 DB 297 LKLIARLEFFEDRTERTNPNMFRIGNLVLYL-----IITHMNAIYFALISKEIFGT 351
 QY 580 DSRIGLWNLGDIQKPYNSGGLGSPSIKDKYVATYALFTFSSLSLVGFGVNSPTNSEKI 639
 DB 352 DS---WVYP-----NYSNPEYGLSKRIYISLWSTLTLTIG-ETPPYKDEYL 398
 QY 640 FSTICVMIIGSLMTASTIGNSAITIQLYSTARYHQMVRERFIRHQIIPNPLRQLEE 699
 DB 399 FVVIDELVGLVIALIATIGNSMISNMNASRAEFQAKIDSIKQYMKRKYTKDLETRVIR 458
 QY 700 YFQHAMSYTNGIDIMNANVLKGFPELQADICLHNRSLLOHCKPFRGATKCLALAMKFK 759
 DB 459 WEYTLWANKKTVEDEKVLKSLPKLKAELIAINWHLTLRRVRIFQDCEAGLIVELVILKR 518

QY 760 TTHAPGDTLVHAGDLITATYFIRSGSIEILRGDYYV--AIIKNDIFGEPLNTYARPK 817
 DB 519 PAVSPBDYICKKDDIGREMIIEGKLAVAEGITQFVVLGDSGFGEISILNIGSK 578
 QY 818 S---NGDVRLATYCDLHKIHRDLEVLDMYPE 847
 DB 579 SGNRTANIRISIGYSDLFCLSKDLMALTEYPE 612
 RESULT 13
 S74179
 Cyclic nucleotide-gated channel protein - human
 C:Species: Homo sapiens (man)
 C:Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 16-Jul-1999
 C:Accession: S74179
 R:Yu, W.P.; Grunwald, M.E.; Yau, K.W.
 FEBS Lett. 393, 211-215, 1996
 A:Title: Molecular cloning, functional expression and chromosomal localization of a h
 A:Reference number: S74179; MUID:96409310; PMID:8814292
 A:Accession: S74179
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-695 <YUW>
 A:Experimental source: retina
 C:Genetics:
 A:Map position: 2
 C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleoti
 C:Keywords: ion channel; ion transport; membrane protein
 F:482-607/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CA
 Query Match 7.18; Score 433; DB 2; Length 695;
 Best Local Similarity 25.68; Pred. No. 1.7e-19;
 Matches 129; Conservative 103; Mismatches 204; Indels 68; Gaps 17;
 QY 362 KIKERTNHTVEKYQVLSIGADVL-PEYKLAQRIHRWTLIHSFFRAVMDLILLVY 420
 DB 138 KCNTNTSNTEEEKTKRKDAIYVDPSSNL--YFWMIAIALP--VFYMWYLL--IC 188
 QY 421 TAVETPYSAAPFLKETEGRPATECGYACOPLAVVDLIVIMEYVILINFRITYVANE 480
 DB 189 RACDELQSEYLM-----WLVLDYSADVLVLDVLRVRGFELEQG- 230
 QY 481 EVVSHPRIVHY-FKGMFLDMVAIIPDL--LIFSGSEELIGLKTARILRVAR 537
 DB 231 LMSVSDTRIMQHYTTQFQKLDVLSVPTDLAYLKVGTNPE-----VRFNRLKFSLE 286
 QY 538 KLDRTS-----EYGAVLELLMCTFALIAHMLACIWAIGMEQPHMDSRIQWLN 588
 DB 287 FFDRTERTNYPNMFRIQNLVLYL-----IITHMNAIYFALISKEIFGTDS--WVY- 337
 QY 589 LGDQIGKPYNSGGLGSPSIKDKYVATYALFTFSSLSLVGFGVNSPTNSEKIFSCVMLIG 648
 DB 338 --PNISSIPENG-----RLSKRIYISLWSTLTLTIG-ETPPYKDEYLIVVVDLVG 388
 QY 649 SIMVASIFGNVAIQLRYSTARYHQMVRERFIRHQIIPNPLRQLEEYFQHAMSYT 708
 DB 389 VLIPIATIGVNGVSMISNMNASRAEFQAKIDSIKQYMKRKYTKDLETRVIRWPFYLMANK 448
 QY 709 NGIDMANVLKGFPELQADICLHNRSLLOHCKPFRGATKCLALAMKFKTTIAPRGDT 768
 DB 449 KTYDEKVELKSLPKLKAELIAINWHLTLKVRIFQDCEAGLIVELVILKRPYFSPGDY 508
 QY 769 LVNAGDLITATYFIRSGSIEILRGDYYV--AIIKNDIFGEPLNTYARPKS---NGDY 822
 DB 509 ICKKGDIGKERYIINBEKLAIVADGVTOFVVLSDGSYFGBISILNIGSKSGNRRYANI 568
 QY 823 RALTVCDLHKIHRDLEVLDMYP 846
 DB 569 RSIGYSDLPCLSKDLMALTEYF 592
 RESULT 14
 S55349

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 10:36:34 ; Search time 73 Seconds
(without alignments)
2520.056 Million cell updates/sec

Title: US-10-000-151b-3
Perfect score: 6079
Sequence: 1 MPVRRGHVAPONTFLDTIIR.....GOLGALTSQLNRHNSGSPGS 1159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq_19Jun03:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6079	100.0	1159	20	AA1932020
2	6079	100.0	1159	21	AA185405
3	6079	100.0	1159	22	AA185182
4	6079	100.0	1159	22	AA185039
5	6079	100.0	1159	23	AA191697
6	6079	100.0	1159	24	AA185167
7	6079	100.0	1159	24	AA185125
8	6072	99.9	1159	21	AA185406
9	5782	95.1	1163	21	AA183031

10	3493.5	57.5	1177	24	ABP71254	Chimeric ZERG poly
11	3293	54.2	1196	23	AAO14189	Human transporter
12	3241	53.3	1186	24	ABP71253	zebrafish ZERG pol
13	3039.5	50.0	994	23	ABG31253	Human erg2 (h-erg2)
14	3001.5	49.4	958	24	ABG31252	Human erg2 (h-erg2)
15	3001.5	49.4	958	24	ABP72167	Human potassium ch
16	2997.5	49.3	958	23	ABG31251	Human erg2 (h-erg2)
17	2996	49.3	947	24	ABG75824	Transporters and 1
18	2993.5	49.2	958	23	AAO14192	Human transporter
19	2984	49.1	879	23	ABG31254	Human erg2 (h-erg2)
20	2765	45.5	905	23	AAE16789	Human transporter
21	2748.5	45.2	894	24	ABG75828	Transporters and 1
22	2436.5	40.1	888	20	AA171399	Human erg subfam1
23	2203.5	36.2	788	24	ABG75829	Transporters and 1
24	1865.5	30.7	855	22	ABG59695	Drosophila melanog
25	1653	27.2	626	20	AA171398	Human erg subfam1
26	1636	26.9	1017	20	AA122427	Human brain specif
27	1568	25.8	1107	21	AA144907	Human potassium ch
28	1551	25.5	1102	21	AA183028	Rat Elk1 potassium
29	1541.5	25.4	1083	21	AA144779	Human Elk voltage
30	1541.5	25.4	1083	21	AA144780	Human Elk voltage
31	1541.5	25.4	1083	21	AA144781	Human Elk voltage
32	1541.5	25.4	1083	21	AA144782	Human brain specif
33	1540.5	25.3	1083	20	AA144904	Human potassium c
34	1540.5	25.3	1082	21	AA134128	Human potassium ch
35	1536	25.3	1083	21	AA184835	Human alpha-subu
36	1536	25.3	1083	21	AA177738	Human ESK1 (hESK1)
37	1529.5	25.2	1080	21	AA177738	Human potassium ch
38	1515.5	24.9	521	24	ABP54351	Human transporter
39	1511.5	24.9	530	24	ABP54350	Drosophila melanog
40	1485.5	24.4	1311	22	ABB61234	Drosophila melanog
41	1453.5	23.9	1174	22	ABB65304	Human potassium 10
42	1415	23.3	962	21	AA149944	Human potassium 10
43	1402.5	23.1	989	21	AA149945	Human ether a gogo
44	1389.5	23.0	988	22	AAE01043	Human alpha-subu
45	1399.5	23.0	988	22	AA131714	

ALIGNMENTS

RESULT 1	AA1932020	standard; Protein: 1159 AA.
ID	AA1932020;	
XX	AC	AA1932020;
XX	DT	05-JAN-2000 (first entry)
XX	DE	Human cation channel protein.
XX	KW	Cation channel protein; CCP; ion transport; arrhythmia;
XX	KW	diabetes mellitus; seizure; asthma; hypertension; therapy;
XX	KW	protein engineering; human.
XX	OS	Homo sapiens.
XX	FT	Key
XX	FT	Region
XX	FT	Location/Qualifiers
XX	FT	61..119
XX	FT	/note="crystal region"
XX	PD	MO9947923-A2.
XX	PD	23-SEP-1999.
XX	PF	22-MAR-1999;
XX	PF	99WO-US06307.
XX	PR	20-MAR-1998;
XX	PR	98US-0045529.
XX	PR	02-APR-1998;
XX	PR	98US-0054347.
XX	PA	(UYRQ) UNIV ROCKEFELLER.
XX	PI	Mackinnon R;

[illegible]

Db	661	AIIGRLVSGTARYTQMLRVAEFLIRFQIIPRLPQRLREEFQHMWSTNGIDMAVYLGKF	720
Qy	721	PECLQADICLHLNSSLQHCKPRFGATKGCILALAMKFKTTNAPGDTLVHAGDLLTALY	780
Db	721	PECLQADICLHLNSSLQHCKPRFGATKGCILALAMKFKTTNAPGDTLVHAGDLLTALY	780
Qy	781	FISGSIETILGDDVVAVILGKNDIFGEPILNLYARPGKSNQGVRLATYCDLHKIHRDLE	840
Db	781	FISGSIETILGDDVVAVILGKNDIFGEPILNLYARPGKSNQGVRLATYCDLHKIHRDLE	840
Qy	841	VLDMPPEFSDFHWSLETTFNLRDTNMIPGSPGSTELEGGSFRQKRKLSPRRTRDKTE	900
Db	841	VLDMPPEFSDFHWSLETTFNLRDTNMIPGSPGSTELEGGSFRQKRKLSPRRTRDKTE	900
Qy	901	QPEYSAIAGPRRAGRSRRCGPGWGESFSGPSSPESEDEDGPRGSSPLRLVPFS	960
Db	901	QPEYSAIAGPRRAGRSRRCGPGWGESFSGPSSPESEDEDGPRGSSPLRLVPFS	960
Qy	961	PRPGEPPGGEPILAEDEKSSDTQNPISGAFSGVSNIFSPMGDSRGQOYELPCPAPTP	1020
Db	961	PRPGEPPGGEPILAEDEKSSDTQNPISGAFSGVSNIFSPMGDSRGQOYELPCPAPTP	1020
Qy	1021	SLNIPILSSPGRRPRGDESRDLALQRLNLETRLSDMATVYLQLQRMVTPPAYSA	1080
Db	1021	SLNIPILSSPGRRPRGDESRDLALQRLNLETRLSDMATVYLQLQRMVTPPAYSA	1080
Qy	1081	VTPGPGGTSPPLPVPSPPTLRLDSISQVSOQMACBELLPGAPELPOGAPTRRLSLPG	1140
Db	1081	VTPGPGGTSPPLPVPSPPTLRLDSISQVSOQMACBELLPGAPELPOGAPTRRLSLPG	1140
Qy	1141	QLGALTSQPLHRHGSDDPS 1159	
Db	1141	QLGALTSQPLHRHGSDDPS 1159	
RESULT 2			
AAV85405	ID	AAV85405 standard; Protein; 1159 AA.	
XX	AC	AAV85405;	
XX	DT	20-JUN-2000 (first entry)	
XX	DE	Long QT syndrome associated HERG protein.	
XX	KM	HERG; mutation; long QT syndrome; LQT syndrome; gene therapy; human.	
XX	OS	Homo sapiens.	
XX	XX		
XX	XX		
FH	Key	Location/Qualifiers	
FH	Misc-difference	572 /note= "can be substituted with Cys; specifically claimed"	
FT	Misc-difference	588 /note= "can be substituted with Asp; specifically claimed"	
FT	Misc-difference	614 /note= "can be substituted with Val; specifically claimed"	
FT	Misc-difference	630 /note= "can be substituted with Ala; specifically claimed"	
FT			
FT			
XX	WO200006772-A1.		
XX	PD	10-FEB-2000.	
XX	PE	20-JUL-1999; 99WO-US16337.	
XX	PR	27-JUL-1998; 98US-0122847.	
XX	PR	06-JAN-1999; 99US-0226012.	

PA (UTAH) UNIV UTAH RES FOUND.
 XX
 PI Keating MT, Splawski I;
 XX
 DR WPI: 2000-195319/17.
 DR N-PSDB; AAA07601.
 XX
 PT New isolated mutant HERG nucleic acids, useful for developing products
 PT for the diagnosis, prevention and treatment of long QT syndrome
 XX
 PS Claim 11: Fig 8A-B; 163pp; English.
 XX
 CC The invention relates to a HERG protein having a mutation compared to
 CC wild-type HERG, and is useful for developing products for the diagnosis,
 CC prevention and treatment of long QT (LQT) syndrome. The products and
 CC methods can be used for the diagnosis of subjects with LQT syndrome. They
 CC can also be used to screen for drugs for treating or preventing LQT
 CC syndrome. The HERG nucleic acids can also be used for gene therapy and
 CC HERG peptides can be used for peptide therapy. The present sequence
 CC represents the LQT syndrome associated HERG protein. The HERG protein
 CC gene sequence comprising any of the following mutations apart from those
 CC indicated above is also specifically claimed for in the specification.
 CC The mutations arise from specific alterations in the encoding HERG gene
 CC sequence. The mutations can be F29L, N33T, C44Stop, G47V, G53R, R36Q,
 CC C66G, H70R, P72Q, R73frameshift, A78P, A83frameshift, Q81stop, L86R,
 CC P141frameshift, P151frameshift, P241frameshift, V295frameshift, R312C,
 CC P347S, R531Q, L552S, A561T, G584S, W585C, I593T, G604S, D609N, T613M,
 CC L615Y, G626S, P632L, P638E, delK638, M645L, E682stop,
 CC H739frameshift, R752W, T798frameshift, F805S, F805C, R823W, N861L,
 CC K866frameshift, P917L, R920frameshift, R922W, G925frameshift,
 CC P968frameshift, P986frameshift, W1001stop, R1014stop, G1031frameshift,
 CC and P1101frameshift.
 CC
 XX
 SQ Sequence 1159 AA:
 Query Match 100.0%; Score 6079; DB 21; Length 1159;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 481 EVVSHPRIAVHYFKGFLIDMVAIPFDLLIFGSGSEELIGLTKTARLLRVAVARKLD 540
 QY 541 RYSEYGAAYFLMCTFALIAHMLACIWAYIAGNEQPHMDSRGMLNLDQIGKPVNS 600
 DB 541 RYSEYGAAYFLMCTFALIAHMLACIWAYIAGNEQPHMDSRGMLNLDQIGKPVNS 600
 QY 601 GLGSPSIKDKYVVALYTFESSLTSVGCNVSPTNNSKIFISICVMILGSLMYASIFGNVS 660
 DB 601 GLGSPSIKDKYVVALYTFESSLTSVGCNVSPTNNSKIFISICVMILGSLMYASIFGNVS 660
 QY 661 AIIQRLYSTARRHTQMLRVREFTREHQIDNPRLQRLEEFYQHAMSTNGIDMAVLKGF 720
 DB 661 AIIQRLYSTARRHTQMLRVREFTREHQIDNPRLQRLEEFYQHAMSTNGIDMAVLKGF 720
 QY 721 PECLQADICLHLNLSLQHKCPFGAKGCCRALAMFKTTHAPPGDTLVHAGDLTALY 780
 DB 721 PECLQADICLHLNLSLQHKCPFGAKGCCRALAMFKTTHAPPGDTLVHAGDLTALY 780
 QY 781 FISRSTIELRGDVVAIILKNDIFGEPMLLYARPGKSGNDVRLATYCDLHKIHRDDLE 840
 DB 781 FISRSTIELRGDVVAIILKNDIFGEPMLLYARPGKSGNDVRLATYCDLHKIHRDDLE 840
 QY 841 VLDMPYPSDFHFWSSLEITFNLRDTNMIPOSPGSTLELGEFSROKRRKLSFRRTDYTE 900
 DB 841 VLDMPYPSDFHFWSSLEITFNLRDTNMIPOSPGSTLELGEFSROKRRKLSFRRTDYTE 900
 QY 901 QPGEVSAALGPRGAGSSNGRPGGSPGESSGSSPSESEDEGPGRSSPLLVFPSS 960
 DB 901 QPGEVSAALGPRGAGSSNGRPGGSPGESSGSSPSESEDEGPGRSSPLLVFPSS 960
 QY 961 PRPGEPPGGEPLMEDEKESDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
 DB 961 PRPGEPPGGEPLMEDEKESDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
 QY 1021 SLNIPILSSPGRRPGVESRLDALQRLNRLRSLADMAVYQLLQRMWTLVPAYSA 1080
 DB 1021 SLNIPILSSPGRRPGVESRLDALQRLNRLRSLADMAVYQLLQRMWTLVPAYSA 1080
 QY 1081 VTPPGPPTSPPLPVSPPLTLDSLVSQFMACEELPGAPELPOGPTRRSLPG 1140
 DB 1081 VTPPGPPTSPPLPVSPPLTLDSLVSQFMACEELPGAPELPOGPTRRSLPG 1140
 QY 1141 QLGALTSQPLRHGSDPGS 1159
 DB 1141 QLGALTSQPLRHGSDPGS 1159

RESULT 3
 AAM51812
 ID AAM51812 standard; Protein: 1159 AA.
 XX
 AC AAM51812;
 XX
 DT 22-JAN-2002 (first entry)
 XX
 DE Human potassium channel subunit HERG.
 XX
 KW S. cerevisiae; yeast; potassium translocation system; TRK1; TRK2;
 KW TOK1; human erg potassium ion channel; HERG; antiarrhythmic;
 KW antiinflammatory; antidiarrhythmic; HERG modulator; human.
 OS Homo sapiens.
 XX
 PN DEL9953478-AI.
 XX
 PD 11-OCT-2001.
 XX
 PF 06-NOV-1999; 99DE-1053478.
 XX
 PR 06-NOV-1999; 99DE-1053478.
 XX
 PA (LICH/) LICHTENBERG-FRATTE H.
 PA (LUDW/) LUDWIG J.

```

XX  Lichtenberg-Frate H, Ludwig J;
PI
XX  WPI: 2001-603577/69.
DR  N-PSDB: AAI66257.
XX
XX  Genetically modified yeast lacking endogenous potassium transport
PT  activity, useful for identifying e.g. antiarrhythmic agents, includes a
PI  functional human potassium channel
XX
XX  Example 3: Fig 8; 40pp; German.
XX
XX  The present invention relates to a genetically modified Saccharomyces
CC  cerevisiae in which the endogenous potassium-translocation systems (TRK1,
CC  TRK2 and TOK1) are specifically deleted and the human erg potassium ion
CC  channel (HERG) is stably integrated and expressed. This can be used to
CC  identify specific modulators of HERG, which are potentially useful as
CC  antiarrhythmic, antifibrillatory and antiinflammatory agents. The present
CC  sequence is the human HERG protein.
XX
XX  Sequence 1159 AA;
SQ
Query Match 100.0%; Score 6079; DB 22; Length 1159;
Batch Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPVRGHVAPQNTFELDTIRKFEQSRKFIITANAVENCAYICNDGFCELCGSRAEYM 60
DB 1 MPVRGHVAPQNTFELDTIRKFEQSRKFIITANAVENCAYICNDGFCELCGSRAEYM 60
QY 61 QRPCTCDLHGPRTRORRAAQTQAALLGAERKVEIAFYRKDSCFCLVDVVPVKNEGG 120
DB 61 QRPCTCDLHGPRTRORRAAQTQAALLGAERKVEIAFYRKDSCFCLVDVVPVKNEGG 120
QY 121 AVIMFILNFEYVMEKDMGSPADHNTNHRGPTSMIAPGAKTFRLLKLPALLATATARESSV 180
DB 121 AVIMFILNFEYVMEKDMGSPADHNTNHRGPTSMIAPGAKTFRLLKLPALLATATARESSV 180
QY 181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTAMDNHVAGLGAPEERRALVGPSP 240
DB 181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTAMDNHVAGLGAPEERRALVGPSP 240
QY 241 PRSAGQULPSRAHSLNDPASGSSCSLARTSRSCASVRRASSADIDEAMRAGVLPPPP 300
DB 241 PRSAGQULPSRAHSLNDPASGSSCSLARTSRSCASVRRASSADIDEAMRAGVLPPPP 300
QY 301 RHASGAMHPLRSGGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITA 360
DB 301 RHASGAMHPLRSGGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITA 360
QY 361 PIKERTHNTEKVTQVLSGADVLPKYLQAPRIHRWTILHYSFPKAVMDLILLVIY 420
DB 361 PIKERTHNTEKVTQVLSGADVLPKYLQAPRIHRWTILHYSFPKAVMDLILLVIY 420
QY 421 TAVFTPYSAALFKTEEGPPATEGCGYACQPLAVVDLIVDIMEFYDILLINERTTYVANE 480
DB 421 TAVFTPYSAALFKTEEGPPATEGCGYACQPLAVVDLIVDIMEFYDILLINERTTYVANE 480
QY 481 EVVSHPGIAYHYFGWFLIMVAAIPFDLLIFSGSSEELGLLKTARLRIVRARAKLD 540
DB 481 EVVSHPGIAYHYFGWFLIMVAAIPFDLLIFSGSSEELGLLKTARLRIVRARAKLD 540
QY 541 RYSEYGAVALFLMCTFALIAHWLACIYVAGNMOPHMDSRIQWLHNLGDOIGRPYSS 600
DB 541 RYSEYGAVALFLMCTFALIAHWLACIYVAGNMOPHMDSRIQWLHNLGDOIGRPYSS 600
QY 601 GLGGSIDKXVTALYFTFSSLSYVGFNVSPNTNSEKIFSTICVWLIGSLMAYSTFGVNS 660
DB 601 GLGGSIDKXVTALYFTFSSLSYVGFNVSPNTNSEKIFSTICVWLIGSLMAYSTFGVNS 660
QY 661 AIIQRLVSGTARYHTQMLRVREFIRFHOIIPNPLRQRLREYFOHAMSYNGIDMNAVILKGF 720
DB 661 AIIQRLVSGTARYHTQMLRVREFIRFHOIIPNPLRQRLREYFOHAMSYNGIDMNAVILKGF 720

```

```

QY 721 PECIALADICLHLNRSLLQHCKPRFGATKGCULRALAMKERTTHAPPDGLVHAGDLLTALY 780
DB 721 PECIALADICLHLNRSLLQHCKPRFGATKGCULRALAMKERTTHAPPDGLVHAGDLLTALY 780
QY 781 FISRGSIETLRKDVVVALGKNDITGCEPLNLTARPKSGNGVRAITYCDLKHIRHDDLE 840
DB 781 FISRGSIETLRKDVVVALGKNDITGCEPLNLTARPKSGNGVRAITYCDLKHIRHDDLE 840
QY 841 VLDMPERFSDHFWSSLETTFNLROTNMIPGSPSTELGSGFSROKRLSPFRRTDKDTE 900
DB 841 VLDMPERFSDHFWSSLETTFNLROTNMIPGSPSTELGSGFSROKRLSPFRRTDKDTE 900
QY 901 QPGEVSALGPRAGAGPSSRGRPGGPMWESPSSGSPSESDGPGRSSSPLRLVPSS 960
DB 901 QPGEVSALGPRAGAGPSSRGRPGGPMWESPSSGSPSESDGPGRSSSPLRLVPSS 960
QY 961 PRPGEPPGGEPLMDCERKSSDTCNPLSGARSGVSNITSFWDGSKRGROYELPRCPAPTP 1020
DB 961 PRPGEPPGGEPLMDCERKSSDTCNPLSGARSGVSNITSFWDGSKRGROYELPRCPAPTP 1020
QY 1021 SLNTPILSSPGRPRRGDVESRLDALQROLNRLRSLADMATVQLQLOMPLVPAPSA 1080
DB 1021 SLNTPILSSPGRPRRGDVESRLDALQROLNRLRSLADMATVQLQLOMPLVPAPSA 1080
QY 1081 VTPPGPGTSTSPLLPVSPPLTLTDSLQVSOQFVACELPPGAPELPOEGPTRRLSLPG 1140
DB 1081 VTPPGPGTSTSPLLPVSPPLTLTDSLQVSOQFVACELPPGAPELPOEGPTRRLSLPG 1140
QY 1141 QLGALTSQPLHRHSGDPGS 1159
DB 1141 QLGALTSQPLHRHSGDPGS 1159
RESULT 4
AAB70939
ID AAB70939 standard; Protein: 1159 AA.
XX
XX AAB70939;
AC
XX
XX 22-AUG-2001 (first entry)
DE
XX
XX Human eag-related protein HERG.
DE
XX
XX Potassium transporter; K+ transporter; TKHp; mutant; potassium uptake;
KW selectable marker; auxotrophic marker; resistance marker; Trk2p;
KW eag related gene; HERG; potassium ion channel; antiarrhythmic;
KW antifibrillatory; antiinflammatory.
XX
XX Homo sapiens.
OS
XX
XX DEL9941768-A1.
PN
XX
XX 15-MAR-2001.
PD
XX
XX 02-SEP-1999; 99DE-1041768.
PE
XX
XX 02-SEP-1999; 99DE-1041768.
PR
XX
XX (LICH/) LICHTENBERG-FRATE H.
PA
XX
XX Lichtenberg-Frate H;
PI
XX
XX WPI: 2001-246033/26.
DR
XX
XX N-PSDB: AAF61965.
DR
XX
XX New Schizosaccharomyces pombe mutants with defective potassium uptake
PT useful for expressing human eag related gene in screening assays for
PT potassium ion channel modulators -
XX
XX Example 1.4; Page 13; 24pp; German.
XX
XX This invention describes novel Schizosaccharomyces pombe mutants with
CC

```


defective potassium uptake which are obtainable by introducing one or more selectable markers (auxotrophy and/or resistance markers) into the genes for potassium transporter TRKp and/or TRK2p. The invention also describes (1) a genetically modified *S. pombe* strain that expresses human eag related gene (HERG) but does not express endogenous TRKp or TRK2p; (2) a screening assay for specific modulators of HERG potassium ion channel, comprising growing an *S. pombe* strain of type (1) in the presence and absence of a test substance and measuring any change in potassium transport. The mutants can be transformed to express human eag related gene (HERG) and used in screening assays for specific modulators of HERG potassium ion channel, which are potentially useful as antiarrhythmic, antifibrillatory and antiinflammatory agents. This sequence represents the human HERG protein described in the disclosure of the invention.

SQ Sequence 1159 AA;

Query Match	100.0%	Score 6079;	DB 22;	Length 1159;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1159; Conservative	0;	Mismatches	0;	Gaps 0;

QY	1	MPVRGHA	PONTE	EDT	ITIRK	FEQGS	REKIT	IAN	AVENC	AVIT	YCD	NDG	SC	ELG	YS	RAE	V	60	
Db	1	MPVRGHA	PONTE	EDT	ITIRK	FEQGS	SRF	IT	IAN	AVENC	AVIT	YCD	NDG	SC	ELG	YS	RAE	V	60
QY	61	ORPCTC	D	LH	P	R	T	O	R	R	A	A	O	I	A	O	A	L	120
Db	61	ORPCTC	D	LH	P	R	T	O	R	R	A	A	O	I	A	O	A	L	120
QY	121	AVIME	I	L	F	E	V	M	E	K	D	M	V	S	P	A	H	D	180
Db	121	AVIME	I	L	F	E	V	M	E	K	D	M	V	S	P	A	H	D	180
QY	181	RSGG	A	G	A	G	A	G	A	V	V	D	L	T	P	A	P	A	240
Db	181	RSGG	A	G	A	G	A	G	A	V	V	D	L	T	P	A	P	A	240
QY	241	PRSA	P	G	O	L	P	S	R	A	H	S	L	N	D	A	S	G	300
Db	241	PRSA	P	G	O	L	P	S	R	A	H	S	L	N	D	A	S	G	300
QY	301	RHAST	G	A	M	H	P	L	R	S	G	L	N	S	T	S	D	S	360
Db	301	RHAST	G	A	M	H	P	L	R	S	G	L	N	S	T	S	D	S	360
QY	361	PKIK	E	R	T	N	N	E	K	Y	O	V	I	S	L	A	D	V	420
Db	361	PKIK	E	R	T	N	N	E	K	Y	O	V	I	S	L	A	D	V	420
QY	421	TAVF	P	Y	S	A	A	F	L	K	T	E	E	G	P	A	T	E	480
Db	421	TAVF	P	Y	S	A	A	F	L	K	T	E	E	G	P	A	T	E	480
QY	481	EVS	H	P	G	R	I	A	V	H	E	F	G	M	F	L	D	M	540
Db	481	EVS	H	P	G	R	I	A	V	H	E	F	G	M	F	L	D	M	540
QY	541	RYSE	G	A	V	E	L	M	C	T	I	A	I	C	I	A	I	C	600
Db	541	RYSE	G	A	V	E	L	M	C	T	I	A	I	C	I	A	I	C	600
QY	601	GLG	P	S	I	D	K	V	T	A	L	F	T	S	S	L	T	S	660
Db	601	GLG	P	S	I	D	K	V	T	A	L	F	T	S	S	L	T	S	660
QY	661	AI	O	R	L	S	G	T	A	R	Y	H	T	O	M	L	R	V	720
Db	661	AI	O	R	L	S	G	T	A	R	Y	H	T	O	M	L	R	V	720
QY	721	PE	C	L	A	D	I	C	L	H	N	S	L	O	H	C	K	P	780
Db	721	PE	C	L	A	D	I	C	L	H	N	S	L	O	H	C	K	P	780

QY	781	FISGSGIEILGDDVVAALICKNDITFGEPLNLIYANPGKSGNDVRLTYCDLHKIHRDLE	840
Db	781	FISGSGIEILGDDVVAALICKNDITFGEPLNLIYANPGKSGNDVRLTYCDLHKIHRDLE	840
QY	841	VLDWYPEPSDFHFWSSLETFEFLRDTNNIPGSPGSTELEGGSRORKKLSFRRTDXTDTE	900
Db	841	VLDWYPEPSDFHFWSSLETFEFLRDTNNIPGSPGSTELEGGSRORKKLSFRRTDXTDTE	900
QY	901	QPGVSALGPGRAGAGPSSRNGRPGGPGMGESPSSGSPSSPESEDEGPGRSSPLNLYPSS	960
Db	901	QPGVSALGPGRAGAGPSSRNGRPGGPGMGESPSSGSPSSPESEDEGPGRSSPLNLYPSS	960
QY	961	PRPGEPPGGGPLMEDCEKSSDTCNPLISGAFSGVSNLFSFWGDSRGQYOELPRCPATP	1020
Db	961	PRPGEPPGGGPLMEDCEKSSDTCNPLISGAFSGVSNLFSFWGDSRGQYOELPRCPATP	1020
QY	1021	SLNLIPLSSPPRRRGDVESRLDALOROLNLEFTRLISADNAVTYQLOROMTLVPAYSA	1080
Db	1021	SLNLIPLSSPPRRRGDVESRLDALOROLNLEFTRLISADNAVTYQLOROMTLVPAYSA	1080
QY	1081	VTTGPGPGTSTSPLLPVSPPTLTLDSLSOVSQPMACBELLPPGABELPQEGPTRRLSLPG	1140
Db	1081	VTTGPGPGTSTSPLLPVSPPTLTLDSLSOVSQPMACBELLPPGABELPQEGPTRRLSLPG	1140
QY	1141	QLGALITSQLHRHGSDEGS	1159
Db	1141	QLGALITSQLHRHGSDEGS	1159

RESULT 5
AAU99167

AC AAU99167;

DT 24-SEP-2002 (first entry)

Human ether-a-go-go related protein, HERG.

Human; human ether-a-go-go related gene; HERG; KCRI1;

KW long quaternarion; LVI; single nucleotide polymorphism; calcium allylamine; potassium channel.

yy
05 Homo sapiens.

XX WO200242735-A2
PN

XX 30-MAY-2002
PD

XX
DE 30-0000-2001.

XX 30-0000-200000-21434000

XX

CHAPTER TWO AND EIGHT

XX

XX

XX

PT comprises contacting a compound with a potassium channel and rat

FI CEDEBETAL COMA LIBRARY (ACNOL) POLYPEPTIDES, and determining activity XX

FS
XX

claim of, page 100-102; 104pp; English:

biochemical activity of a potassium channel

CC cloned from a rat cerebellar cDNA library (NCRI), and determining the
CC compound with a structure comprising a PC polypeptide and a polypeptide
CC activity of the PC polypeptide in the presence and absence of the
CC compound, where a difference in the activities indicates modulation of
CC biological activity of PC. Also include are identifying (M2) a candidate
CC compound that modulates the biological activity of a complex comprising a

human ether-a-go-go-related gene (HERG) channel polypeptide and a KCRI polypeptide, identifying (M3) a candidate compound as a modulator of KCRI expression, modulating (M4) PC function in a subject, comprising administering to the subject a substance that provides expression of a KCRI-encoding nucleic acid molecule in a cell or tissue, where modulated PC function is desired, screening (M5) for susceptibility to a drug-induced cardiac arrhythmia in a subject, comprising obtaining a biological sample from the subject and detecting a polymorphism of a KCRI gene in the biological sample from the subject, where the presence of the polymorphism indicates the susceptibility of the subject to a drug-induced cardiac arrhythmia, an oligonucleotide pair, where a first oligonucleotide of the pair hybridises to a first portion of a KCRI gene which includes a polymorphism of the KCRI gene, and the second oligonucleotide of the pair hybridises to a second portion of the KCRI gene that is adjacent to the first portion and a set of antisense oligonucleotide primers, suitable for amplifying a portion of a KCRI gene which includes a polymorphism of the KCRI gene. (M1) is useful for identifying a compound that modulates biological activity of PC, especially HERG, for modulating PC function (i.e. modulating HERG activity) in a mammal, by preparing a composition comprising the compound and administering the composition. The compound is useful for treating or preventing long QT syndrome (LQT) and is useful in drug designing. The present sequence represents human HERG.

Sequence 1159 AA;

Query Match 100.0%; Score 6079; DB 23; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MPVRGHAAPQNTFLDTIIRKFEQSKFTIANANAVENCAYIYCDGCELCGYSRAEVM 60
DB 1 MPVRGHAAPQNTFLDTIIRKFEQSKFTIANANAVENCAYIYCDGCELCGYSRAEVM 60
OY 61 ORPCTCDLHGPRTORAAOIAOALGAERKVIATYRKDGSCFLTVDPVPAKNDG 120
DB 61 ORPCTCDLHGPRTORAAOIAOALGAERKVIATYRKDGSCFLTVDPVPAKNDG 120
OY 121 AVIMEILNFEVMEKDMVGSAPADTNHNGPPTSWLAPRAKTFRLKLPALLATARESSV 180
DB 121 AVIMEILNFEVMEKDMVGSAPADTNHNGPPTSWLAPRAKTFRLKLPALLATARESSV 180
OY 181 RSGGAGGAGAGVAVVDLTPAPSSSLADEVTANDNVAGICGPEERRALVGPSP 240
DB 181 RSGGAGGAGAGVAVVDLTPAPSSSLADEVTANDNVAGICGPEERRALVGPSP 240
OY 241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCSAVRRASSADIDEARAGVLPSP 300
DB 241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCSAVRRASSADIDEARAGVLPSP 300
OY 301 RHASTGAMHPLRSGSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360
DB 301 RHASTGAMHPLRSGSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360
OY 361 PKIKERTHNTEKTVQVLSLADVLPEKLAQRHRTIILHSPFKAWMMLILLIY 420
DB 361 PKIKERTHNTEKTVQVLSLADVLPEKLAQRHRTIILHSPFKAWMMLILLIY 420
OY 421 TAVFTPYSAALFKETEGEPATEGACQPLAVVDLVDIMETIYDILINFRITYVANNE 480
DB 421 TAVFTPYSAALFKETEGEPATEGACQPLAVVDLVDIMETIYDILINFRITYVANNE 480
OY 481 EVVSHPGRIAVHYFKGMFLIDMVAIPFDLLIFGSGSEELIGLKTARLLRVYARKLD 540
DB 481 EVVSHPGRIAVHYFKGMFLIDMVAIPFDLLIFGSGSEELIGLKTARLLRVYARKLD 540
OY 541 RYSEYGAIVFLMLCTFALLIHMILACTIYVAILGNMOPRMDSRIGLMLGQIGKPYSS 600
DB 541 RYSEYGAIVFLMLCTFALLIHMILACTIYVAILGNMOPRMDSRIGLMLGQIGKPYSS 600
OY 601 GLGSPSTIKDKYVLTALYFTFSSLTSGFGNVSPNTNSEKIFISICVWLIGSLMYASTFGNVS 660
DB 601 GLGSPSTIKDKYVLTALYFTFSSLTSGFGNVSPNTNSEKIFISICVWLIGSLMYASTFGNVS 660

```

```

OY 661 AIIQRYSGTARYHTOMLRVEEFIRHQIPNPLRQRLREYFOHASTYNGIDMNAVLKGF 720
DB 661 AIIQRYSGTARYHTOMLRVEEFIRHQIPNPLRQRLREYFOHASTYNGIDMNAVLKGF 720
OY 721 PECLADICLHNSRLDCHCKPFRGATYGCRLAMKFKTTTHAPPGDTLVHAGDILLIY 780
DB 721 PECLADICLHNSRLDCHCKPFRGATYGCRLAMKFKTTTHAPPGDTLVHAGDILLIY 780
OY 781 FISRSSTIILGDDVVVATLIGKNDIFGEPGLNIYARPKSGNGVRLTYCDLKHIRDDILE 840
DB 781 FISRSSTIILGDDVVVATLIGKNDIFGEPGLNIYARPKSGNGVRLTYCDLKHIRDDILE 840
OY 841 VLDMPERSDHFWSLETFNLRDNTMIPGSPSTELGCFSRQKRKLSEFRRTDKPTE 900
DB 841 VLDMPERSDHFWSLETFNLRDNTMIPGSPSTELGCFSRQKRKLSEFRRTDKPTE 900
OY 901 QPGEVSALGPBAGAGPSSRRPGGPGWGESPPSGSPSESSDEGPGRSSPLRLVPSS 960
DB 901 QPGEVSALGPBAGAGPSSRRPGGPGWGESPPSGSPSESSDEGPGRSSPLRLVPSS 960
OY 961 PRPEGEPPGPIEMDCCKSSDTCNPLSGATSVNIFSPWGDSDRGROYELPCPAPT 1020
DB 961 PRPEGEPPGPIEMDCCKSSDTCNPLSGATSVNIFSPWGDSDRGROYELPCPAPT 1020
OY 1021 SLNLIPLSSPGRRPRGDVESRLDALQRLNLETRLSDAMATVQLQRLQMTLVPPAYSA 1080
DB 1021 SLNLIPLSSPGRRPRGDVESRLDALQRLNLETRLSDAMATVQLQRLQMTLVPPAYSA 1080
OY 1081 VTTGPGGTSTSPLLPVSPPTLTIDLSQVSGFACCELPAGBELPQEGPTRLSLPG 1140
DB 1081 VTTGPGGTSTSPLLPVSPPTLTIDLSQVSGFACCELPAGBELPQEGPTRLSLPG 1140
OY 1141 QLGALTSOPLHRHGSPPGS 1159
DB 1141 QLGALTSOPLHRHGSPPGS 1159

```

```

RESULT 6
ABP71697
ID ABP71697 standard; Protein: 1159 AA.
XX
AC ABP71697;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human HERG potassium channel polypeptide.
XX
KW HERG; cardiotoxicity; cardiac arrhythmia; ether-a-go-go-related gene;
KW potassium channel; astemizole; human.
OS Homo sapiens.
PN WO2003006988-A1.
XX
PD 23-JAN-2003.
XX
PF 02-JUL-2002; 2002WO-EP07364.
XX
PR 13-JUL-2001; 2001EP-0202689.
XX
PA (JUNC ) JANSSEN PHARM NV.
XX
PI Heylen GICM, Janssen CGM, Jurzak M, Van Assouw HPM;
XX
DR WPT: 2003-267961/26.
XX
DR N-PSDB: ABZ/6227.
XX
PT Screening test compounds to induce cardiotoxicity/cardiac arrhythmia,
PT by incubating human ether-a-go-go-related gene channel, with a
PT reference and test compound, and measuring effect of the test on
PT reference compound -
XX

```

PS Claim 17; Page 27-29; 43pp; English.
 XX
 CC The invention relates to an assay for screening test compounds for their
 CC capability to induce cardiotoxicity or cardiac arrhythmia in a subject.
 CC The method involves incubating a source containing human ether-a-go-go-
 CC related gene channel (HERG) or a fragment of it with a reference compound
 CC and a test compound, and measuring the effect of the test compound on the
 CC amount of reference compound bound to HERG. A radiolabeled astemizole is
 CC useful as the reference compound in the method for screening test
 CC compounds for their capability to induce cardiotoxicity or cardiac
 CC arrhythmia in a subject. The present sequence represents a human
 CC HERG potassium channel polypeptide.
 XX
 XX
 SQ Sequence 1159 AA;
 Query Match 100.0%; Score 6079; DB 24; Length 1159;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPVRGHVAPONTFLDTIIRKEGOSRKFITANARVENCAYICNDPFCELGYSRAEYV 60
 DB 1 MPVRGHVAPONTFLDTIIRKEGOSRKFITANARVENCAYICNDPFCELGYSRAEYV 60
 QY 61 QRPCTDFLHGPTORRAAQAIALGAEBKVEIAFTRKDGSCFLIVDVVPKNEDEG 120
 DB 61 QRPCTDFLHGPTORRAAQAIALGAEBKVEIAFTRKDGSCFLIVDVVPKNEDEG 120
 QY 121 AVIMFLNFEVMEKMGSPADHTNRGPTSMALPGRAKPTRLPALALTARESSV 180
 DB 121 AVIMFLNFEVMEKMGSPADHTNRGPTSMALPGRAKPTRLPALALTARESSV 180
 QY 181 RSGGAGGAGAPGAVVVDVLTTPAAPSSSESLALDEVYAMDNHVAAGLPAERERLAVPGSP 240
 DB 181 RSGGAGGAGAPGAVVVDVLTTPAAPSSSESLALDEVYAMDNHVAAGLPAERERLAVPGSP 240
 QY 241 PRSAPQQLPSRAHSINPDASSGSCSLAATRSRESCASVRRASSADIDEAMRAGVLP 300
 DB 241 PRSAPQQLPSRAHSINPDASSGSCSLAATRSRESCASVRRASSADIDEAMRAGVLP 300
 QY 301 RRASTGAMHPLRSGLINSTSDSLVYRTISKIPQITLNPVDLKGDFLASPSDEBITA 360
 DB 301 RRASTGAMHPLRSGLINSTSDSLVYRTISKIPQITLNPVDLKGDFLASPSDEBITA 360
 QY 361 PKIKERTHNVTEKVTQVLSLADVLEPYLQAPRIHRMTILHSPFKAWMDLILLVLY 420
 DB 361 PKIKERTHNVTEKVTQVLSLADVLEPYLQAPRIHRMTILHSPFKAWMDLILLVLY 420
 QY 421 TAVFTPYSAFLIKETEGRPAATECGYACQPLAVVDLIVDIMEVDILINFRTYYVANE 480
 DB 421 TAVFTPYSAFLIKETEGRPAATECGYACQPLAVVDLIVDIMEVDILINFRTYYVANE 480
 QY 481 EYVSHHGRTAVHYFKGMFLIDWVAATPFDLITFGSSSEELIGLTKTARLLRYVARAKID 540
 DB 481 EYVSHHGRTAVHYFKGMFLIDWVAATPFDLITFGSSSEELIGLTKTARLLRYVARAKID 540
 QY 541 RYSEYGAALVFLMCFALIAHMLACTIWAIGMEQPHMDSRGMHNLGDOIGKRYNS 600
 DB 541 RYSEYGAALVFLMCFALIAHMLACTIWAIGMEQPHMDSRGMHNLGDOIGKRYNS 600
 QY 601 GLGGSIKDKYVYALYFTFSSLTSGVGNVSPNTSEKIFSIICVMLIGSLMAYASIFGNV 660
 DB 601 GLGGSIKDKYVYALYFTFSSLTSGVGNVSPNTSEKIFSIICVMLIGSLMAYASIFGNV 660
 QY 661 ATIORLYSGTARVHOMLVREIRPHOINPRLQRLLEYFQHAMSTYTGIDMNAVLKGF 720
 DB 661 ATIORLYSGTARVHOMLVREIRPHOINPRLQRLLEYFQHAMSTYTGIDMNAVLKGF 720
 QY 721 PECLQADICLHLNRSILQCKPFGATKGLRALAKFKTTNAPRPDITVHAGDILLTALY 780
 DB 721 PECLQADICLHLNRSILQCKPFGATKGLRALAKFKTTNAPRPDITVHAGDILLTALY 780
 QY 781 FLSRGSIEILLRGDVVAAILGKNDIFGEPLNLVAPGKSGNDVRLTYCDLHKITHRDDLE 840
 DB 781 FLSRGSIEILLRGDVVAAILGKNDIFGEPLNLVAPGKSGNDVRLTYCDLHKITHRDDLE 840

DB 781 FLSRGSIEILLRGDVVAAILGKNDIFGEPLNLVAPGKSGNDVRLTYCDLHKITHRDDLE 840
 QY 841 VLDMYEFSDFHFWSSLEITFNLTNMGSPGSTELEGGFSQRKRKLSFRRTDKDTE 900
 DB 841 VLDMYEFSDFHFWSSLEITFNLTNMGSPGSTELEGGFSQRKRKLSFRRTDKDTE 900
 QY 901 QPGEVALGPRGAGAPSSRGPRGPGWGESPPSSPESSEDEGGRSSPLRLVPFSS 960
 DB 901 QPGEVALGPRGAGAPSSRGPRGPGWGESPPSSPESSEDEGGRSSPLRLVPFSS 960
 QY 961 PRPGEPGGEPLMECEKSSDPCNLSAFAFSVSNIFPSWGSRRQVQELPRCAPTP 1020
 DB 961 PRPGEPGGEPLMECEKSSDPCNLSAFAFSVSNIFPSWGSRRQVQELPRCAPTP 1020
 QY 1021 SLINILPSSPGRRPRDVEDSRDALROLNRLTETRISADMAVYLOLLOROMTLVPPAYSA 1080
 DB 1021 SLINILPSSPGRRPRDVEDSRDALROLNRLTETRISADMAVYLOLLOROMTLVPPAYSA 1080
 QY 1081 VTTPGPGPTSTPLPLVSPPLTLTDLISQVQFMACEELPPGAPELPQGPTRRLSLPG 1140
 DB 1081 VTTPGPGPTSTPLPLVSPPLTLTDLISQVQFMACEELPPGAPELPQGPTRRLSLPG 1140
 QY 1141 QLGALTSQPLHRHSDPGS 1159
 DB 1141 QLGALTSQPLHRHSDPGS 1159
 RESULT 7
 ABP71255
 ID ABP71255 standard; Protein; 1159 AA.
 XX
 AC ABP71255;
 XX
 DT 28-APR-2003 (first entry)
 XX
 DE Human HERG polypeptide.
 XX
 KW Teleost; zebrafish; ZERG; cardiovascular; antiarrhythmic; cyostatic;
 XX neuroprotective; gene therapy; potassium channel; human; HERG.
 OS Brachydanio rerio.
 PN WO2003006502-A2.
 XX
 PD 23-JAN-2003.
 XX
 PF 11-JUL-2002; 2002MO-IB04280.
 XX
 PR 13-JUL-2001; 2001US-305396P.
 XX
 PA (ARTE-) ARTEMIS PHARM GMBH.
 XX
 PI Langheirrich U;
 XX
 DR WPI: 2003-210421/20.
 XX
 PT New teleost (specifically zebrafish) ERG genes, which encode ERG family
 PT potassium channels, useful for studying e.g. cardiac or brain function,
 PT or for developing treatments for e.g. cardiac diseases, cancer or
 PT neurological diseases -
 XX
 PS Examples; Page 50-55; 55pp; English.
 XX
 CC The invention relates to a teleost (zebrafish) ERG (ZERG) polypeptide.
 CC The ZERG nucleic acid molecule is useful for studying cardiac function,
 CC abnormal heart beat phenotype, or long QT syndrome (an abnormality of
 CC cardiac muscle repolarization that predisposes affected individuals to
 CC lethal arrhythmias). The zebrafish comprising ZERG gene is useful as
 CC models for cardiac function or disease. The ZERG genes are particularly
 CC useful for in (non-)cardiac researches, or for developing treatments for
 CC cardiac diseases, tumours or cancers, brain and nervous system disorders
 CC or neurological diseases, or insulin-related diseases. The present
 CC sequence represents a human HERG (ether-a-go-go-related) polypeptide.

XX Sequence 1159 AA:
 Query Match 100.0%; Score 6079; DB 24; Length 1159;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVRGHVAPONTFLDITIRKFEQSRKFIITANARVENCAYICNDGFCGLCYSRAEVM 60
 DB 1 MPVRGHVAPONTFLDITIRKFEQSRKFIITANARVENCAYICNDGFCGLCYSRAEVM 60

QY 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYVVPVKKNDG 120
 DB 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYVVPVKKNDG 120

QY 121 AVIMFIINFVYMEKDMWGSFAHDTNHRGPTSMIAPGRAKTFRLKLPALLALTARESSV 180
 DB 121 AVIMFIINFVYMEKDMWGSFAHDTNHRGPTSMIAPGRAKTFRLKLPALLALTARESSV 180

QY 121 AVIMFIINFVYMEKDMWGSFAHDTNHRGPTSMIAPGRAKTFRLKLPALLALTARESSV 180
 DB 121 AVIMFIINFVYMEKDMWGSFAHDTNHRGPTSMIAPGRAKTFRLKLPALLALTARESSV 180

QY 181 RSGGAGAGAGAVVVDVLTLPAPSSSESLADEVTAMDNHVAGLGAPEERRALVPGSP 240
 DB 181 RSGGAGAGAGAVVVDVLTLPAPSSSESLADEVTAMDNHVAGLGAPEERRALVPGSP 240

QY 241 PRSAPGQLSPRAHSINPDASGSSCSLARTSRSCASVRAASADIEAMRAGVLP 300
 DB 241 PRSAPGQLSPRAHSINPDASGSSCSLARTSRSCASVRAASADIEAMRAGVLP 300

QY 301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360
 DB 301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360

QY 361 PKIKERHNTVEKYTOVLSIGADVLPEYKIQAPRIHRMTILHYSPEKAVMDLILLVITY 420
 DB 361 PKIKERHNTVEKYTOVLSIGADVLPEYKIQAPRIHRMTILHYSPEKAVMDLILLVITY 420

QY 421 TAVTPTSAFLLKETEGPRATECGYACQPLAVVDLVDIMFTVDLILNRTTYVANE 480
 DB 421 TAVTPTSAFLLKETEGPRATECGYACQPLAVVDLVDIMFTVDLILNRTTYVANE 480

QY 481 EVVSHPRIAVHYKGFLLIDMVAIIPFDLIFGSSSEELIGLKTARLLRVYARKLD 540
 DB 481 EVVSHPRIAVHYKGFLLIDMVAIIPFDLIFGSSSEELIGLKTARLLRVYARKLD 540

QY 541 RYSEYGAUVFLMCTALLAHMLACTIWAIGNEQPHMDRIGMLHMGDIQKPYNS 600
 DB 541 RYSEYGAUVFLMCTALLAHMLACTIWAIGNEQPHMDRIGMLHMGDIQKPYNS 600

QY 601 GLGSPSTKDKYVTLVTFESSLTSGVGNVSPNTNSKIFISICVMLIGSLMAYASIFGNVS 660
 DB 601 GLGSPSTKDKYVTLVTFESSLTSGVGNVSPNTNSKIFISICVMLIGSLMAYASIFGNVS 660

QY 661 AIIORLSTGARYTOMLRVEEFIRFQIOPNPLRQRLSEYFQAHMSTNGIDMAVAKGF 720
 DB 661 AIIORLSTGARYTOMLRVEEFIRFQIOPNPLRQRLSEYFQAHMSTNGIDMAVAKGF 720

QY 721 PECLQADICHLNLSLQHCKPRFGATKGCILALAMFKTTHADPGDTLVHAGDLTALY 780
 DB 721 PECLQADICHLNLSLQHCKPRFGATKGCILALAMFKTTHADPGDTLVHAGDLTALY 780

QY 781 FISGSEILIRGVDVVAIIGKNDIFGEPNLNYPARKSNGVYALTYCDLKHIRDDLE 840
 DB 781 FISGSEILIRGVDVVAIIGKNDIFGEPNLNYPARKSNGVYALTYCDLKHIRDDLE 840

QY 841 VLDNYPFESDHFWMSLITFNLRTNMIPGSPSTELGEGFSRORRKLFRRTDQTE 900
 DB 841 VLDNYPFESDHFWMSLITFNLRTNMIPGSPSTELGEGFSRORRKLFRRTDQTE 900

QY 901 QPGEVSALGPRAGAGPSSRGPRGPGWESBSPSPSESEDEGPGRSSSPLRLVFPSS 960
 DB 901 QPGEVSALGPRAGAGPSSRGPRGPGWESBSPSPSESEDEGPGRSSSPLRLVFPSS 960

QY 961 PRPGEPPGGEPLMEDCEKSSDTGNPLSGAFSGVSNITFSWGDGRGOYQELPRCPAPTP 1020
 DB 961 PRPGEPPGGEPLMEDCEKSSDTGNPLSGAFSGVSNITFSWGDGRGOYQELPRCPAPTP 1020

DB 961 PRPGEPPGGEPLMEDCEKSSDTGNPLSGAFSGVSNITFSWGDGRGOYQELPRCPAPTP 1020
 QY 1021 SLNLIPLSSPGRRPGRDVESRLDALQROLNRLERLSADMATVQLQROMTLVPPAYSA 1080
 DB 1021 SLNLIPLSSPGRRPGRDVESRLDALQROLNRLERLSADMATVQLQROMTLVPPAYSA 1080

QY 1081 VTPPGPGPTSTSPILLPVSPPLTLTLDLSISQVSQFMACEELPPGAPBELPQEGPTRLSLPG 1140
 DB 1081 VTPPGPGPTSTSPILLPVSPPLTLTLDLSISQVSQFMACEELPPGAPBELPQEGPTRLSLPG 1140

QY 1141 QLGALTSQPLHRHGSDDPS 1159
 DB 1141 QLGALTSQPLHRHGSDDPS 1159

RESULT 8
 AAY85406
 ID AAY85406 standard; Protein: 1159 AA.
 XX AAY85406;
 AC AAY85406;
 DT 19-JUN-2000 (first entry)
 DE Long QT syndrome associated HERG protein.
 KW HERG; mutation; long QT syndrome; LQT syndrome; gene therapy; human.
 OS Homo sapiens.
 PN W0200006772-A1.
 XX 10-FEB-2000.
 PD 20-JUL-1999; 99WO-US16337.
 PF 27-JUL-1998; 98US-0122847.
 PR 06-JAN-1999; 99US-0226012.
 PA (UTAH) UNIV UTAH RES FOUND.
 XX Keating MT, Splawski I;
 PI WPI: 2000-195319/17.
 DR N-PSDB; AAA07602.
 PT New isolated mutant HERG nucleic acids, useful for developing products for the diagnosis, prevention and treatment of long QT syndrome -
 XX Disclosure: Fig 8A-B; 163pp; English.
 CC The invention relates to a HERG protein having a mutation compared to wild-type HERG, and is useful for developing products for the diagnosis, prevention and treatment of long QT (LQT) syndrome. The products and methods can be used for the diagnosis of subjects with LQT syndrome. They can also be used to screen for drugs for treating or preventing LQT syndrome. The HERG nucleic acids can also be used for gene therapy and HERG peptides can be used for peptide therapy. The present sequence represents the LQT syndrome associated HERG protein.

QY 99.9%; Score 6072; DB 21; Length 1159;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPVRGHVAPONTFLDITIRKFEQSRKFIITANARVENCAYICNDGFCGLCYSRAEVM 60
 DB 1 MPVRGHVAPONTFLDITIRKFEQSRKFIITANARVENCAYICNDGFCGLCYSRAEVM 60

QY 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYVVPVKKNDG 120
 DB 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYVVPVKKNDG 120

```

QY 121 AVIMFLNEVNEKDMVSGPAHDTHNRGPTSMALPGRAKTERLKLALLATARESSV 180
    |||
Db 121 AVIMFLNEVNEKDMVSGPAHDTHNRGPTSMALPGRAKTERLKLALLATARESSV 180
QY 181 RSGGAGAGAGAVVVDVDTLPAPASSESLALDEVTAAMDNHVAGLGPABERRALVGPSP 240
    |||
Db 181 RSGGAGAGAGAVVVDVDTLPAPASSESLALDEVTAAMDNHVAGLGPABERRALVGPSP 240
QY 241 PRGAPOLPSRAHSLNPDASGSSCLARPRRESGASVRRASADDIEMRAGVLP PPP 300
    |||
Db 241 PRGAPOLPSRAHSLNPDASGSSCLARPRRESGASVRRASADDIEMRAGVLP PPP 300
QY 301 RAHSTGAMHPLRSGLNSTSDSLVRYRTSKIPQITLNVVDKGFPLSPSDREITIA 360
    |||
Db 301 RAHSTGAMHPLRSGLNSTSDSLVRYRTSKIPQITLNVVDKGFPLSPSDREITIA 360
QY 361 PKIKERTHNVEKVTQVLSGADVLPDEYKLAQAPRIHRMTLHYSPFKAVMDMLILLVIY 420
    |||
Db 361 PKIKERTHNVEKVTQVLSGADVLPDEYKLAQAPRIHRMTLHYSPFKAVMDMLILLVIY 420
QY 421 TAVFTPYSAFLAKETEEGPATCEGYACQPLAVVDLIVDMFTVDLINFRTTYVANE 480
    |||
Db 421 TAVFTPYSAFLAKETEEGPATCEGYACQPLAVVDLIVDMFTVDLINFRTTYVANE 480
QY 481 EVVSHPGRIAVHYFKGMFLIDMYAATPFDDLIFGSGSEELIGLKTARLLRLVAVARKLD 540
    |||
Db 481 EVVSHPGRIAVHYFKGMFLIDMYAATPFDDLIFGSGSEELIGLKTARLLRLVAVARKLD 540
QY 541 RYSEYGAANVLELLMCTFALIAHMLACIWAIGNMEQPHMDSRIGMNLHLDQIGKPYNS 600
    |||
Db 541 RYSEYGAANVLELLMCTFALIAHMLACIWAIGNMEQPHMDSRIGMNLHLDQIGKPYNS 600
QY 601 GLGGRSTKDKYVYALFTFSSLSVYRGVNSPTNSKITSICVMLIGSLMYASIGVNS 660
    |||
Db 601 GLGGRSTKDKYVYALFTFSSLSVYRGVNSPTNSKITSICVMLIGSLMYASIGVNS 660
QY 661 AIIOQLYSGTARYTQMLRVREFIRFHQIPNPLRQLEEFYOHAMSTYNGIDNNAYLKG 720
    |||
Db 661 AIIOQLYSGTARYTQMLRVREFIRFHQIPNPLRQLEEFYOHAMSTYNGIDNNAYLKG 720
QY 721 PECLQADICLHLNRSLLQHCPPRGATKGLRALAMKFKTTTHAPPGDVLVHAGDLTALY 780
    |||
Db 721 PECLQADICLHLNRSLLQHCPPRGATKGLRALAMKFKTTTHAPPGDVLVHAGDLTALY 780
QY 781 F1RSGSTEILRGVYVAAILGKNDIFGEPLLYARPGKSNQDVALTYCDLHKIHRDLE 840
    |||
Db 781 F1RSGSTEILRGVYVAAILGKNDIFGEPLLYARPGKSNQDVALTYCDLHKIHRDLE 840
QY 841 VLDMPPEFSDHFWSSLEITFNLDJTNMIPGSPSTEEGGSQORRKLIFRRRTDKTE 900
    |||
Db 841 VLDMPPEFSDHFWSSLEITFNLDJTNMIPGSPSTEEGGSQORRKLIFRRRTDKTE 900
QY 901 QPEVSALGPGAGAGPSSRGPRGPGWGESPPSSGESSEDEGPGRSSPRLVFPSS 960
    |||
Db 901 QPEVSALGPGAGAGPSSRGPRGPGWGESPPSSGESSEDEGPGRSSPRLVFPSS 960
QY 961 PRPPGEPGGEPLMEDECKSSDTCNPLSGAFSGVSNFISFWGDSRGQYDELPCPAPTP 1020
    |||
Db 961 PRPPGEPGGEPLMEDECKSSDTCNPLSGAFSGVSNFISFWGDSRGQYDELPCPAPTP 1020
QY 1021 SLNINPLSPGRRPRGVDRESRLDALQRLNLETRLSADAMATVQLLQOROMTIVPPAYSA 1080
    |||
Db 1021 SLNINPLSPGRRPRGVDRESRLDALQRLNLETRLSADAMATVQLLQOROMTIVPPAYSA 1080
QY 1081 VTPPGGPTSTSPLLPVSPILTLDLSQVSOFMACEELPQGAPELLQEGPFRRLSLPG 1140
    |||
Db 1081 VTPPGGPTSTSPLLPVSPILTLDLSQVSOFMACEELPQGAPELLQEGPFRRLSLPG 1140
QY 1141 QLGALTSQPLHRHSGDPGS 1159
    |||
Db 1141 QLGALTSQPLHRHSGDPGS 1159

```

```

RESULT 9
AAV83031
ID AAV83031 standard; Protein; 1163 AA.
XX
AC AAV83031;
XX
DT 04-JUL-2000 (first entry)
XX
DE Rat Erg1 potassium channel membrane protein.
XX
KW Elk1; elk2; eag1; eag2; erg1; potassium channel; membrane protein;
KW drug screening; hypertension; renal failure; diabetes insipidus;
KW diabetic nephropathy; hypothyroidism; goiter; hypoparathyroidism;
KW pancreatic insufficiency; diabetes mellitus; cystic fibrosis;
KW salivary; salivary insufficiency; diabetes mellitus; current flow;
KW ion flux; transcription; signal transduction; assay; detection;
KW rat.
XX
OS Rattus rattus.
XX
PN W0200012546-A1.
XX
PD 09-MAR-2000.
XX
PE 31-AUG-1999; 99WO-US19902.
XX
PR 31-AUG-1998; 98US-0098413.
XX
PA (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX
PI McKinnon D, Dixon JE;
XX
DR WPI: 2000-256585/22.
XX
PT Novel mammalian potassium channel genes and polypeptides encoded by
PT them for screening drugs useful for treating diseases such as
PT hypertension, acute renal failure, diabetes insipidus and
PT hypothyroidism -
XX
PS Disclosure; Page 90-95; 102pp; English.
XX
CC Potassium channel genes e.g. elk1, elk2 or eag2 are useful for
CC identifying modulators which are useful for treating hypertension,
CC acute and chronic renal failure, diabetes insipidus, diabetic
CC nephropathy, hypothyroidism, goiter, hyperparathyroidism,
CC pancreatic insufficiency, diabetes mellitus, cystic fibrosis,
CC salivary, salivary insufficiency. The availability of the gene
CC sequences provides a tool for research into the physiological
CC characteristics of the various genes and proteins for potassium
CC channels including the development of medicines effective for
CC treating disease conditions associated with mutations or defects in
CC potassium channels and the screening of drugs to ensure that
CC potassium channels are not blocked or physiologically affected by
CC those drugs. The channel proteins encoded by these genes are also
CC useful to measure changes in potassium concentration, membrane
CC potential, current flow, ion flux, transcription, signal
CC transduction, receptor-ligand interaction and second messenger
CC concentrations. See also GENESEQ records AA29334-293336.
XX
XX
Sequence 1163 AA:
Query Match 95.1%; Score 5782; DB 21; Length 1163;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1111; Conservative 12; Mismatches 36; Indels 4; Gaps 2;
QY 1 MPVARGHVAPONFLODTIIRKFEQOSRKFTIANARVENCAVIYCNDFCELGYSRAEVM 60
    |||
Db 1 MPVARGHVAPONFLODTIIRKFEQOSRKFTIANARVENCAVIYCNDFCELGYSRAEVM 60
QY 61 QRPCTCDFLHGPRTORRAAAQIAQALLGAERKEVIAFYRRDGSCLCLVDVVPVKNEDG 120
    |||
Db 61 QRPCTCDFLHGPRTORRAAAQIAQALLGAERKEVIAFYRRDGSCLCLVDVVPVKNEDG 120

```

```

QY 121 AVIMEILNEFYVMEKDMGSPAHDTNHRGPTSWLAPGRAKTFRLKLPALLATARESSV 180
DB 121 AVIMEILNEFYVMEKDMGSPAHDTNHRGPTSWLAPGRAKTFRLKLPALLATARESSPM 180
QY 181 RSGGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAIGPAEERATVGP--G 238
DB 181 RTGSGTSPGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAIGPAEERATVGPASA 240
QY 239 SPSPAPGOLPSPRAHSLNPDAGSSCSLARTRESGASVARRASADDIEAMRAGVLP 298
DB 241 SPVASIPGPHSPRAQNSLNPDASSCSLARTRESGASVARRASADDIEAMRAGALPL 300
QY 299 PPRHASTGAMHPLRSGGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREI 358
DB 301 PPRHASTGAMHPLRSGGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREI 360
QY 359 IAPKIKERTHNTEVTOVLSLADVDLPEYKLOARIRHWTILHSPPKAVWDMLILLY 418
DB 361 IAPKIKERTHNTEVTOVLSLADVDLPEYKLOARIRHWTILHSPPKAVWDMLILLY 420
QY 419 IYTAVFTEPYSAAFILKETEEGPATECGYACOPLAIVVDLIVDIMEFVILINFRTYYNA 478
DB 421 IYTAVFTEPYSAAFILKETEDGSOAPDCGYACOPLAIVVDLIVDIMEFVILINFRTYYNA 480
QY 479 NEEVYSHRGRIAVHFKGMFLIDMVAALPFDLLIFGSGSEELIGLKTAARLRLTVARAK 538
DB 481 NEEVYSHRGRIAVHFKGMFLIDMVAALPFDLLIFGSGSEELIGLKTAARLRLTVARAK 540
QY 539 LDRYSEYGAAYVELLMCTFEALIAHMLACIWAIGMBOPHMDSRIGMHNGDOIGKPYN 598
DB 541 LDRYSEYGAAYVELLMCTFEALIAHMLACIWAIGMBOPHMDSRIGMHNGDOIGKPYN 600
QY 599 SSGLGGPSIKDKYVALYEFSSLSVGRGVNSPPTNSEKIFPSICVMLIGSLMVASIRGN 658
DB 601 SSGLGGPSIKDKYVALYEFSSLSVGRGVNSPPTNSEKIFPSICVMLIGSLMVASIRGN 660
QY 659 VSAIIQRIYSGTARHTOMLRREFIRRHQIIPNPRQRLSEYFOHAWSTYNGIDMNAVYK 718
DB 661 VSAIIQRIYSGTARHTOMLRREFIRRHQIIPNPRQRLSEYFOHAWSTYNGIDMNAVYK 720
QY 719 GPPECLADICLHLNRSLLOHCKPRGATKGLRALAMKFKTTHAPRGGTLVHAGDILTA 778
DB 721 GPPECLADICLHLNRSLLOHCKPRGATKGLRALAMKFKTTHAPRGGTLVHAGDILTA 780
QY 779 LYFISRGSIELLRGDVVAALGKNDIFGEPLNLVYARPGKSGNDVYALTYCDLHKIHRDL 838
DB 781 LYFISRGSIELLRGDVVAALGKNDIFGEPLNLVYARPGKSGNDVYALTYCDLHKIHRDL 840
QY 839 LEVIDMYPEFSDHFWSSLEITFNLRDTNMIPGSPGSTELEGFSQORRKLISFRRTTKD 898
DB 841 LEVIDMYPEFSDHFWSSLEITFNLRDTNMIPGSPGSTELEGFSQORRKLISFRRTTKD 900
QY 899 TEOPGEVVAL--GPRAGAGPSSRGPRGPGWGESPPSSPSESSEDDGPGSSSPPLV 956
DB 901 TEQGEVVALGOGPARVPGPSCRGQPGWGESSSPSPSESSEDDGPGRSSSPPLV 960
QY 957 PFSSPRPGEPPGGEPLMEDDEKSSDTCNPLSGAFSGVSNIFSEWGDGRGOYOLPRCP 1016
DB 961 PFSSPRPGEPPGGEPLMEDDEKSSDTCNPLSGAFSGVSNIFSEWGDGRGOYOLPRCP 1020
QY 1017 APTBSLNLPLSSSGRRRGVESHDLAQLRQLNLETRLASADATVYLQILQOROMTLVP 1076
DB 1021 AAPASLNLPLSSSGRRRGVESHDLAQLRQDNLETRLASADATVYLQILQOROMTLVP 1080
QY 1077 AYSAVTTPGPGPTSNPLPYSPPLTLLDLSLSOVSOMACEBELPGAPBELPOGEPTRRL 1136
DB 1081 AYSAVTTPGPGPTSNPLPYSPPLTLLDLSLSOVSQFVAPELPAAGAPLPODEPTRRL 1140
QY 1137 SLPGOLGALTSQLHRHGSDDPGS 1159
DB 1141 SLPGOLGALTSQLHRHGSDDPGS 1163

```

```

RESULT 10
ABP71254
ID ABP71254 standard; Protein; 1177 AA.
XX
AC ABP71254;
XX
DE 28-APR-2003 (first entry)
XX
DE Chimeric ZERG polypeptide.
XX
KW Teleost; zebrafish; ZERG; cardiovascular; antiarrhythmic; cytostatic;
KW neuropeptide; gene therapy; potassium channel; chimeric..
XX
OS Brachydanio rerio.
XX
OS Homo sapiens.
XX
PN WO2003006502-A2.
XX
PD 23-JAN-2003.
XX
PE 11-JUL-2002; 2002WO-IB04280.
XX
PR 13-JUL-2001; 2001US-305396P.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Langheirlich U;
XX
DR WPI, 2003-210421/20.
XX
PT New teleost (specifically zebrafish) ERG genes, which encode ERG family
PT potassium channels, useful for studying e.g. cardiac or brain function,
PT or for developing treatments for e.g. cardiac diseases, cancer or
PT neurological diseases -
XX
PS Claim 24; Page 46-50; 55pp; English:
XX
CC The invention relates to a teleost (zebrafish) ERG (ZERG) polypeptide.
CC The ZERG nucleic acid molecule is useful for studying cardiac function,
CC abnormal heart beat phenotype, or long QT syndrome (an abnormality of
CC cardiac muscle repolarization that predisposes affected individuals to
CC lethal arrhythmias). The zebrafish comprising ZERG gene is useful as
CC models for cardiac function or disease. The ZERG gene are particularly
CC useful for in (non-)cardiac researches, or for developing treatments for
CC cardiac diseases, tumours or cancers, brain and nervous system disorders
CC or neurological diseases, or insulin-related diseases. The present
CC sequence represents a chimeric zebrafish ZERG polypeptide.
XX
SQ Sequence 1177 AA;

```

Query Match 57.5%; Score 3493.5; DB 24; Length 1177;
Best Local Similarity 61.1%; Pred. No. 4.1e-275;
Matches 757; Conservative 107; Mismatches 227; Indels 147; Gaps 32;

```

QY 1 MPVRRGHVAPONTFLDTIIRKFEQSGKFFITANARVENCAYIYNDGFCGLGYSRAEYV 60
DB 1 MPVRRGHVAPONTFLDTIIRKFEQSGKFFITANARVENCAYIYNDGFCGLGYSRAEYV 60
QY 61 QRCPCDPLHPRQGRRAAQLAOLLAGAEERKVEIAYRRDGSQFCLVNVVPKKNDS 120
DB 61 QRCPCDPLHPRQGRRAAQLAOLLAGAEERKVEIAYRRDGSQFCLVNVVPKKNDS 120
QY 61 QRCPCDPLHPRQGRRAAQLAOLLAGAEERKVEIAYRRDGSQFCLVNVVPKKNDS 120
DB 61 QRCPCDPLHPRQGRRAAQLAOLLAGAEERKVEIAYRRDGSQFCLVNVVPKKNDS 120
QY 121 AVIMEILNEFYVMEKDMV--GSPAHDTNHRGPTSWLAPGRAKTFRLKLPALLATARES 178
DB 121 AVIMEILNEFYVMEKDMV--GSPAHDTNHRGPTSWLAPGRAKTFRLKLPALLATARES 178
QY 179 SVRSAGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAIGLP---AEERATL 234
DB 179 SVRSAGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAIGLP---AEERATL 234
QY 173 SLPKDHFEG-----VVVDY-----LQPSHEVALKLDL-----MSRDSCLKSEYQAL 214
DB 235 V--GPGSPRAAPGOLPSPRAHSLNPDAGSSCSLARTRESGASVARRASADDIEAMR 292

```



```

Db 215 IQQTPSSCELSPPSPNSDLRLEPSGP-----LKHSHSRSMHSLRRASSLHDDGNR 267
OY 293 AGVLPPEPRHASTGAMHPLRSGLLNSTSDILVRYRTISKIQIQTENFDLKGDPPLASP 352
Db 268 -----DOMSDLKPSNLNSTSSDLMKRRHTIGRIPOVYISF---GSDRLRPP 310
OY 353 T-SDRRIIAP-KIKERTHWYTER---VTVLSIGADVLEPYKLOAPRIHRTWILHSPRK 407
Db 311 SPTTEIILAPSKIKORSQNVSEKVTQVYVLSIGADVLEPYKLOAPRIHRTWILHSPRK 370
OY 408 AWMDMILLVLYTAVFTYSAFLIKETEEGPATECGYACQPLAVDLIVDIMEYDI 467
Db 371 AWMDMILLVLYTAVFTYSAFLIKETEEGPATECGYACQPLAVDLIVDIMEYDI 430
OY 468 LINFRTTYNANEVYSHPGRIAVHFKGMFLIDWAAIPDLLIFGSGSEELIGLKTA 527
Db 431 LINFRTTYNANEVYSHPGRIAVHFKGMFLIDWAAIPDLLIFGSGSEELIGLKTA 490
OY 528 RLRLRYVARKLDRYSEYGAVALFLIMCTFALIAHMLACIWAIGMEQPHMDSRIGWLH 587
Db 491 RLRLRYVARKLDRYSEYGAVALFLIMCTFALIAHMLACIWAIGMEQPHMDSRIGWLH 550
OY 588 NIGDOI GKYPNSSGJGSPKIKKYVYALYFTFSSLTSGVGFVNSPNTSEKIFISICVMI 647
Db 551 NIGDOI GKYPNSSGJGSPKIKKYVYALYFTFSSLTSGVGFVNSPNTSEKIFISICVMI 610
OY 648 GSIMYASTIFGNYSATIIORLYSGTARYHOMLRVERIRRHQIPNPLROLEEFYFOANMY 707
Db 611 GSIMYASTIFGNYSATIIORLYSGTARYHOMLRVERIRRHQIPNPLROLEEFYFOANMY 670
OY 708 TNGIDMNAVALKGFPECLQADICLHNRSLLOQCKAFRGASKCGLRLAARFRTIHAAPD 767
Db 671 TNGIDMNAVALKGFPECLQADICLHNRSLLOQCKAFRGASKCGLRLAARFRTIHAAPD 730
OY 768 TIYHAGDILLTALYFYSRSGIETLRGDVVAIIIGKNDIFGEPLNLTARPEKSGNDVRLTY 827
Db 731 TIYHAGDILLTALYFYSRSGIETLRGDVVAIIIGKNDIFGEPLNLTARPEKSGNDVRLTY 790
OY 828 CDLHKIHRDLEVIDMVEPESDHFWSLEITFNLRDTMMIGSPGSTLEEGFSGRQR 887
Db 791 CDLHKIHRDLEVIDMVEPESDHFWSLEITFNLRDTMMIGSPGSTLEEGFSGRQR 850
OY 888 KLSFR-----RTDKDTEQGEVSALCPGRAGAPSSRGPGGPGWGESPSG-PSPE 939
Db 851 RMLPRNRNPDGMDRGMOTYVQPCSPGVNHR-GAIPISQ-----WDLCDGSPASIS 903
OY 940 SSEDGP---GRS---SSPLRLVPPSSPPRPGEPGEPGLMEDCEK-SSDTCNPLSGAFS 992
Db 904 SSEDGP---GRS---SSPLRLVPPSSPPRPGEPGEPGLMEDCEK-SSDTCNPLSGAFS 955
OY 993 -----GVSNISFWDGSRGROYELPRCAPTPSLNIPLSPPGRPRPGVDSRDLAQ 1046
Db 956 YTAAPLNTISGVSYLSDRASEYSEQRSSSAVOACYHHSPCVGRP-NOLOARLELQ 1014
OY 1047 ROLNLETLASDMAVLOLLOROMTLVPAYSAVT-----TPGPGPTS-----T 1091
Db 1015 SOLNLETLASDMAVLOLLOROMTLVPAYSAVT-----TPGPGPTS-----T 1074
OY 1092 SPULPV---SPULPITLDSLSQVSO-----FMACEELPPGAPF---LPQEGP-- 1132
Db 1075 TPTSLITDASDPKSPKSDVSLSEKSPDSSSGLIHVLVASTDTMSMPTETELSVSPGPLL 1134
OY 1133 -----TRRLSLPGOL---GALTSOP-LHRHGSOP 1157
Db 1133 QPGLILCSSLRFPSPDLSEGFOTLEGSEIQRHNSDP 1172

```

```

RESULT 11
AA014189
ID AA014189 standard: protein; 1196 AA.
AC AC
XX AA014189;
XX

```

```

DT 03-MAY-2002 (first entry)
XX
DE Human transporter and ion channel TRICH-6.
XX
KW Human; transporter and ion channel; TRICH; transport disorder;
KW neurological disorder; muscle disorder; immunological disorder;
KW cell proliferative disorder; neuroprotective; noctropic;
KW cerebrotective; immunosuppressive; cytostatic; respiratory; muscular;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN W0200204520-A2.
XX
PD 17-JAN-2002.
XX
PE 05-JUL-2001; 2001WO-US21448.
XX
PR 07-JUL-2000; 2000US-216547P.
PR 14-JUL-2000; 2000US-218232P.
PR 21-JUL-2000; 2000US-220112P.
PR 28-JUL-2000; 2000US-221839P.
XX
PA (INCY-) INCYTE GENOMICS INC.
PI Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;
PI Burford N, Greene BD, Sanjaunvala MS, Baughn MR, Yao MG, Yang J;
PI Patterson C, Gandhi AR, Harlilla AJA, Tribouley CM, Walla NK;
PI Au-Young J, Walsh RT, Ramkumar J, Lu Y, Lu DAM, Azimtal Y, Lal P;
PI Elliott VS, Nguyen DB, Xu Y, Sellhammer JJ, Borowsky ML, Khan FA;
PI Kearney L, Thangavelu K, Das D, Policky JL,
XX
XX MPI: 2002-205969/26.
DR N-PSDB: AAL44672.
XX
PT New human transporters and ion channel polypeptides for diagnosing,
PT treating or preventing transport, neurological, muscle, immunological
PT and cell proliferative disorders -
XX
PS Claim 1; Page 151-153; 230pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human transporter and ion channel proteins, designated
CC TRICH-1-TRICH-32. The sequences can be used in the treatment of
CC transport, neurological, muscle, immunological and cell proliferative
CC disorders. The present sequence is a protein of the invention.
XX
SQ Sequence 1196 AA:

```

```

Query Match 54.2%; Score 3293; DB 23; Length 1196;
Best Local Similarity 55.2%; Pred. No. 9.3e-259;
Matches 705; Conservative 126; Mismatches 241; Indels 206; Gaps 25;

```


QY	295	VLP--P--PRAS-----TGAMHPLRSLGINTSDSDVRYKRTISKIPDITLNFVULK	344
Db	284	VHKNIFEDRNASDNGNVRKGPNNHKSSLLSGTSDBMLNKYSTINKIPDITLNFSEVK	343
QY	345	GDPEFLAS-PTSDRELIAPKIKERTHNHYEKTQVLSIGADVLPREKLOAPRHNHTIAY	403
Db	344	TEKNSSPSSDDKTIAPKYVDKRHNHYEKTQVLSIGADVLPREKLOAPRHNHTIAY	403
QY	404	SPFAVMDWMLLLVITYAVETPYSAAEFLKETEGRPATECGYACOPLAVDIVDIMF	463
Db	404	SPFAVMDWMLLLVITYAIFTPYSAAEFLNDREB-QKRECGYSCSPLANVDIVDIMF	462
QY	464	IYDILINPRTYVANAENEVSHPRINHYKGFILDMVAAPDILIPSSGSE---L	520
Db	463	IYDILINPRTYVANAENEVSDPAKHAIHYKGFILDMVAAPDILIPSSGSETTTL	522
QY	521	IGLKTARLRLVYVARKLDTRYSEYGAVLEFLMCTFEALIAHWACIYVAIGNEDOPAMD	580
Db	523	IGLKTARLRLVYVARKLDTRYSEYGAVALMLMCIIFALINHWAICITYALGNVERPLYT	582
QY	581	SRICMLHNGDOIGKPYN-SSGLGSPSIKDKYVYALYFTSSSLTSVGFVGNVSPNTSEKI	639
Db	583	DKIGWMLSDIGOOIKRYNDSOSSGSPSIKDKYVYALYFTSSSLTSVGFVGNVSPNTSEKI	642
QY	640	FSICVMLIGSLMAYASIEGNVSAIIIOIRLYSGARHTQWLRYREFRHOINPRLORLEE	699
Db	643	FSICVMLIGSLMAYASIEGNVSAIIIOIRLTSGARHNMOLRYKEFLRFHQINPRLORLEE	702
QY	700	YFOHAMSTYNGIDMNAVALKGFPECLQADICHLNRSLLOHCKPFGAKGCLRALAMFK	759
Db	703	YFOHAMSTYNGIDMNMVYLKGFPECLQADICHLNQTLIOCKAKFGAKGCLRALAMFK	762
QY	760	TTNHPRGDTLVHAGDLELTALYFISRGSEIIRGQVYVALILKNDIFGEPRLYLXAPCKSN	819
Db	763	TTNHPRGDTLVHAGDVTALYFISRGSEIILKDIYVALILKNDIFGEMVYLXAPCKSN	822
QY	820	GDVALATYCDLHKIHRDLEVLDMYPEFSDFHWSLIEITFNLRDNTN---IPGSPGSTE	876
Db	823	ADVALATYCDLHKIQRDLEVLDMYPEFSDFHFLTNLELTJNLNRHESKAKDLLRQSMND	882
QY	877	LEGFSRORRKRKLSFRRRRTDKDTEOPGEVSALGPRAGAGFSSKRGPGGWPGESSPGSPS	936
Db	883	SEGNCKLRKRKLSFESEGEK-----NSTN	908
QY	937	SPESSED-----EGPGSSSPLTLVPFSSPRPPGPGGRLMDECKSSDT	983
Db	909	DPEBSADTIRHYOSSKRHFEEKKSRSSSFISIT-----DDE	944
QY	984	CNPLSGAFSGVSNIFSEWDSRGROYOELAPCPAPTPSLNIPL-----	1027
Db	945	QKPL--FSGIVDSSPGIGKASGIDFEE---TYPTSGRMHIDKRSBCKDITDMRSER	997
QY	1028	-----SSPGRRPR-----GDVESRLDALQROLNRLLETRLSDAMATVL	1064
Db	998	ENAHPOPEDESSPASAORAAMGISETESDILTAGEVEORIDILLOEOINLRLESOMTIDITQTL	1057
QY	1065	QLLOROMTIVPAYSAVTPPGPGTSTPL-----PVSGPIPT-LTLDSLSOVSOMA	1116
Db	1058	QLLORQTTIVPAYSAWTA---GSBYORPITQLMRTSOPKASIKTDNFSFSSQOPEFLD	1114
QY	1117	CEELPGAPBELPOEG-----PTRLSLP-G	1140
Db	1115	LEKSLKSKESLSSGVHLNTASENDNLTSLEKQDSDLSLELRLQRKTYVYHIRHSLPDS	1174
QY	1141	QLGALTSOPLRHRSDDG	1158
Db	1175	SLSTVGVIGLRRHVSDPG	1192

RESULT 12
ABP71253
ID ABP71253 standard; Protein; 1186 AA..
XX

AC	ABP11253;
XX	
DT	28-APR-2003 (first entry)
XX	
DE	zebrafish ZERG polypeptide.
XX	
KW	Telostei; zebrafish; ZERG; cardiovascular; antiarrhythmic; cytostatic;
KM	neuroprotective; gene therapy; potassium channel.
OS	
XX	Brachydanio rerio.
XX	
PN	MO2003006502-A2.
PD	
PF	23-JAN-2003.
XX	
PR	11-JUL-2002; 2002WO-IB04280.
XX	
PA	13-JUL-2001; 2001US-305396P.
XX	
XX	(ARTE-) ARTEMIS PHARM GMBH.
PI	
XX	Langheinrich U;
DR	
N-PSDB:	WPt: 2003-210421/20.
XX	
PT	New teleost (specifically zebrafish) ERG genes, which encode ERG family
PT	potassium channels, useful for studying e.g. cardiac or brain function,
PT	or for developing treatments for e.g. cardiac diseases, cancer or
PT	neurological diseases -
XX	
PS	Examples; Page 40-45; 55pp; English.
CC	
CC	The invention relates to a teleost (zebrafish) ERG (ZERG) polypeptide.
CC	The ZERG nucleic acid molecule is useful for studying cardiac function,
CC	abnormal heart beat phenotype, or long QT syndrome (an abnormality of
CC	cardiac muscle repolarization that predisposes affected individuals to
CC	lethal arrhythmias). The zebrafish comprising ZERG gene is useful as
CC	models for cardiac function or disease. The ZERG genes are particularly
CC	useful for in (non-)cardiac researches, or for developing treatments for
CC	cardiac diseases, tumours or cancers, brain and nervous system disorders
CC	or neurological diseases, or insulin-related diseased. The present
CC	sequence represents a zebrafish ZERG polypeptide, an ERG family potassium
CC	channel polypeptide.
SQ	
	Sequence 1186 AA;
	Query Match 53.3%; Score 3241; DB 24; Length 1186;
	Best Local Similarity 57.5%; Pred. No. 1,6e-254;
	Matches 717; Conservative 121; Mismatches 252; Indels 158; Gaps 36
OY	1 MPVRBGHAPOMTFLDITIRKPEGSQRKITINAVENCNAVITYCNGFELGYSRAEYW 60
DB	1 MEVRBGHALQNTTYDITIRKFDGQRKLINAOKNCGIILYCNGFCOMFESEAEIM 60
OY	61 QRPCTCDLHGERTORRAAQAIALGAEEKVEIAFYRKDGSCFLCVDDVVPVKNEDG 120
DB	61 QQSCTCFELYGVTGKTMSALGOALAQLSGEEKVELLYXSKGTCTRCPLDIVPVKNESG 120
OY	121 AVIMEILNEEVMEKDMV--GSPAHDTNHGRPPTSWLADGRAKTFRLKLPAIALATARES 178
DB	121 VVIMEILNQELLDPDSMKKGGLKORMAN-----SWLRAGQRRRMHLRMPSTL-RVKROP 172
OY	179 SVRSSGAGAGAGAPGVVDVDTLPAAFPSESIALADEVYTAMDHNHVAGLG---AEERRAL 234
DB	173 SLPKDHFE-----VVVDY----LQPSHEEVALKKQ-----MSPDCSLKSTOAL 214
OY	235 V--GGSPRSAPRGOLPSPRAHSLLNPDAAGSSCSLARTRSRSCASVRRASSADDIEARK 292
DB	215 IQQTSSCSCLSPPPSPDRLEPSCP-----LKHSISRRESMHLRRASSLIHDIGNR 267
OY	293 AGVLPPPRHASTGAMHPLRSGILNSTSDSIVLVRYRTIKSIQITLNFLVDLKGDPFLASP 352
	:


```

Db      165 --GPG-----PG-----
QY      294 GVLPPPHASTGAMHPLRSLGSLNSTSDSLVRYRTISKIPQITLNFVNLKGDPLASPT 353
Db      170 -----TGR-----GKRTISQIQFTLNFVNLKGRSSST 201
QY      354 SDREIAP-KIKERTHNTKVTQVLSIGADVLPKYQADRIHRTILHSPKAVDW 412
Db      202 TEIILAPHKVTERTQNTKVTQVLSIGADVLPKYQADRIHRTILHSPKAVDW 261
QY      413 LILLVITYTAFTPYSAFLKETEPRATECGAACPLAVNDLIYIMTIDLINER 472
Db      262 LILLVITYTAFTPYSAFLKETEPRATECGAACPLAVNDLIYIMTIDLINER 320
QY      473 TTYVNAEEVSHBGRIVAHYFKGMFLIDMAAIPFDLLIFGSGSEE--LIGLKTARL 529
Db      321 TTYVNTNEVYSHRRIRIAVHYFKGMFLIDMAAIPFDLLIFRTGSEDTTLIGLKTARL 380
QY      530 LRLVAVARKIDRYSEYGAAYFLMLCTFALIAHMLACIWAIGNMEQPHMSRIGMLHNL 589
Db      381 LRLVAVARKIDRYSEYGAAYFLMLCTFALIAHMLACIWAIGNMEQPHMSRIGMLHNL 440
QY      590 GDOIGKPNSSG-LGGPSIKDKYVTALETSSLSLVGFGVNSPTNSEKIFSTCYMLIG 648
Db      441 GVQLGKRTNGSDPASGSPVDKYVTALETSSLSLVGFGVNSPTNSEKIFSTCYMLIG 500
QY      649 SLMTASIFGNVSAIIOQLRYSGTARYHTQMLRVREFIRFHQIPNPLRQLEEFYQWAMS 708
Db      501 SLMTASIFGNVSAIIOQLRYSGTARYHTQMLRVREFIRFHQIPNPLRQLEEFYQWAMS 560
QY      709 NGIDMNAVLYKFPPELQADICLHNRSLQHCPRGATKCLRALANKFYTTHAPGDT 768
Db      561 NGIDMNAVLYKFPPELQADICLHNRSLQHCPRGATKCLRALANKFYTTHAPGDT 620
QY      769 LVHAGDLTALTYFISRGSEITELRGDVYVAIIGKNDIEPEINIXAPKSGDVALNYC 828
Db      621 LVHAGDLTALTYFISRGSEITELRGDVYVAIIGKNDIEPEINIXAPKSGDVALNYC 680
QY      829 DLHRIHRDLEVLDMYEPFSDHFWSSLEITFNLRDT-----NMIPGSPGSTELEGFSR 883
Db      681 DLHRIHRDLEVLDMYEPFSDHFWSSLEITFNLRDT-----NMIPGSPGSTELEGFSR 738
QY      884 QRRKRLSRRRTDKDTEQPEVSALGPRAGAPSSRRPRGPGWGESSSGPS--SPSSS 941
Db      739 -----LSDNQGSPHE--LGP-----QFSPKGYSLGPGSQ 767
QY      942 EDEGPGRSSSPRLRVFPSSPPRPGEPGEPLEMEDECKESDTCNPLSGAFGVSNIFFFW 1001
Db      768 NSMGAG-----PCAFGHPDAAPL-----SISDASGLW 795
QY      1002 GDSRGROYQELPCRPAPTPSLNIPPLSSGRRPRGDVE-----SRDALQQRINLET 1054
Db      796 PE-----LLOEMP-----PRHSP-QSPQEDPCWPLKGLSRLQOQAQNRLES 838
QY      1055 RLSDMATVQLQLOROM-----TLVPKATSAVYTTTPRG-----PIST 1091
Db      839 RVSSDLSTIQLQDKPMQGHASYTLEAPASNDLALVPID-SETPSPRLPQGLFLPPAQ 897
QY      1092 SP-----LLVPSP-----LPTLT--DSLQVSOQMACEBLPPG-----APELQOE 1130
Db      898 TPSTGDDLDCCPKHRNSSPRKPHLAVAMDKTLAPSE--QOEDEGLMPLASLAPLE 953
QY      1131 -----GPTRLSLPGOLGALTQO-PLHRGSDPG 1158
Db      954 VQGLICGFCFS-SLPEHLGVSVPKQLDPRHGSDDPG 987

```

```

RESULT 14
ABG31252
ID ABG31252 standard; Protein; 958 AA.
XX AC
XX ABG31252;

```

```

DT      21-OCT-2002 (first entry)
XX
DE      Human erg2 (h-erg2) protein variant.
XX
KM      Human; erg2; hypotensive; hypertensive; cytostatic;
KM      antileukemia; nephrotropic; potassium channel inhibitor;
KM      hypertension; hypertensive; renal failure; benign prostatic hyperplasia;
KM      prostate cancer; infertility; potassium channel.
XX
OS      Homo sapiens.
XX
PN      WO200242417-A2.
XX
PD      30-MAY-2002.
XX
XX      16-NOV-2001; 2001WO-US43490.
XX      20-NOV-2000; 2000US-249981P.
PR      (MERI ) MERCK & CO INC.
XX
PA      Folander KL, McKenna EJ, Swanson RJ, Liu Y;
XX      WPL; 2002-583376/62.
XX      N-PSDB; ABR88232.
DR
XX
XX      New isolated human-erg2 potassium channel subunit, useful for treatment
XX      of hypertension, hypotension, renal failure, benign prostate
XX      hyperplasia, prostate cancer and infertility
XX
XX      Claim 8; Fig 2B; 53pp; English.
XX
PS      This invention relates to an isolated human erg2 potassium channel
PS      subunit protein. The erg2 protein of the invention is useful for
PS      identifying activators or inhibitors of potassium channels containing
PS      the protein. The erg2 protein is also useful in counter screens for
PS      assays designed to identify activators and inhibitors of other drug
PS      targets. The protein is useful for treating hypotension, hypertension,
PS      renal failure, benign prostatic hyperplasia, prostate cancer, and
PS      infertility. The activators and inhibitors of potassium channels
PS      containing h-erg2 protein, identified using this protein are useful for
PS      treating or preventing conditions as described above, where the
PS      activity of potassium channels containing h-erg2 protein is abnormal.
PS      The nucleic acid encoding the human erg2 protein is useful in
PS      various diagnostic methods, and a DNA or RNA oligonucleotide probe
PS      is useful in diagnostic methods to identify patients having
PS      variant forms of h-erg2 gene, to determine the level of expression
PS      of RNA encoding h-erg2, or to isolate genes homologous to h-erg2 from
PS      other species. The DNA sequence is also useful in gene therapy
PS      techniques to introduce the h-erg2 protein into cells of the target
PS      organs. The present sequence represents the human erg2 protein variant
PS      of the invention.
SQ
Sequence 958 AA:
Query Match 49.4%; Score 3001.5; DB 23; Length 958;
Best local similarity 53.7%; Pred. No. 3.8e-235;
Matches 659; Conservative 83; Mismatches 139; Indels 347; Gaps 30;
QY      1 MPVRGHVAPONTFLDTIRFEQSRKFTIANRVCMAVYICNDGFCGLGSRAYM 60
Db      1 MPVRGHVAPONTFLDTIRFEQSRKFTIANRVCMAVYICNDGFCGLGSRAYM 60
QY      61 QRPCTCDLHPRPQORRAAQAIAALGAERKVEIARYRKGSCFLIVVVPKKNEDG 120
Db      61 QRPCTCDLHPRPQORRAAQAIAALGAERKVEIARYRKGSCFLIVVVPKKNEDG 120
QY      121 AVIMEFLNFVEYMEKDMVGSAPADTNHGRPTSWLAPGRATFKLIPALIA-LTARESS 179
Db      121 AVIMEFLNFVEYMEKDMVGSAPADTNHGRPTSWLAPGRATFKLIPALIA-LTARESS 179
QY      180 VR-----SGAGAGAGAGAVVDVLTLPAPSSSLALDEVYAMDNHVAAGLGRAPERRA 233
Db      180 VR-----SGAGAGAGAGAVVDVLTLPAPSSSLALDEVYAMDNHVAAGLGRAPERRA 233

```

Db	146	QRLLSQSLFGSGSHGRG-----	164
OY	234	LVPGSPSPRASPAGOLPSBRAHSLNDPASGSSCSLARTSRRESACASVRASSADDIEMAMRA	293
Db	165	--GPG-----PG-----	169
OY	294	GVLPPPPRHASGAMHPLRSGILNSTSDSDLYRTYRTISKIPIOTITLNFVDLKCDPLASPT	353
Db	170	-----TGR-----GKYRTISIQIPQFTLNFVEFNLEKRRSSST	201
OY	354	SDREITAP-KIKERTHNTYKTVOLSLGADVLPEYKLOAPRIHMTILHYSFPFAYWDM	412
Db	202	TEIELIAHKVEYERQONTYKTVOLSLGADVLPEYKLOAPRIHMTILHYSFPFAYWDM	261
OY	413	LILLVITYAVFTPYSAAFLLKETEPEGPRATECGACOPLAYVDLIVDMIFVDILINER	472
Db	262	LILLVITYAVFTPYSAAFLLSDQESRGA-CSYTCSPLYVVDLIVDMIFVDLIVNR	320
OY	473	TTYVANEVYSHPERIAVHFPGFLLDMAVAIFPDLLIFGSGSE--LIGLITARL	529
Db	321	TTYVNTNEVYSHPRIRIAVHFEGKMFLLDMAVAIFPDLLIFRTGSDSETTLIGLITARL	380
OY	530	LRLVAVARKLDYSSYGAAVLELLMCTALLAHMLACTIYAGANNEQPMHDSRIGIMHL	589
Db	361	LRLVAVARKLDYSSYGAAVLELLMCTALLAHMLACTIYAGANNEPRLERHKKIMDLSL	440
OY	590	GDQJCKPYNSSG-LGSPSFKDKYVALYFTFSSLSYVGANVSPTNSEKIPISICVMLIG	648
Db	441	GVOLGKRRNGSDPRAGSPVQDKYVALYFTFSSLSYVGANVSPTNSKAYSTICVMLIG	500
OY	649	SLMTASIFGNVSAIIORLYSGTARYHTOMLRVREIRFHQIPNPLRORLEEYFOHAWSYT	708
Db	501	SLMTASIFGNVSAIIORLYSGTARYHTOMLRVKEIRFHQIPNPLRORLEEYFOHAWSYT	560
OY	709	NGIDMNAVLEKFPBELLQADICHLNRSLLQCKPRRGATKGLRALAMKFKTTNAPGDT	768
Db	561	NGIDMNAVLEKFPBELLQADICHLNRRALLOHCPAGSGAGKGLRALAMKFKTTNAPGDT	620
OY	769	LVAHADLLTALYEFIRGSEIELRGGVVAAILGKNPIFGPBLLYARPGKSNQDVALTYC	828
Db	621	LVLHGDVLYSTLYFIRGSEIELRQDVVAAILGKNPIFGPBLLYARPGKSSADVALTYC	680
OY	829	DLKHIRHDDLLEVLDMYEPFSDHEWSSLEITFNLDY-----NMLPGSPGSTELEGFSR	883
Db	681	DLKHIRQADLLEVLDMYPAFAESFMSKLEVTENLMDAAGLHSSPRQAPGSDHQGF--	738
OY	884	QRKRLSRRRRDKTQEQPGEYSALGPGBAGPSSRGAPGPGWGESPPSGSPSPESSED	943
Db	739	-----LSDNOSD-----AAPRLISDASGLW-----PELLQOE	765
OY	944	EGPGRSSPLRLVPPSSPRPGEPPRGEPPLMEDCKSSDTCNPLSGAFSGVNSITSPWCD	1001
Db	766	-----MPRHSPQSPQED--DC-----W--	782
OY	1004	SKRGROYELPRCPAPTEPLLNLPISSPPRRPRGVDRESRLDALOROLNRLLETSLADMATV	1065
Db	783	-----PLK-----LGSRLQEOQAMNRLRESVSDLSRI	811
OY	1064	LQLOLQROM-----TLVPAYSAVTTPO-----GPTST-----	1091
Db	812	LQLOLKPMPOGHASTYLEAPASNDLALPIA-SETTSGPRRUPGFLPPAPGYSGLDD	870
OY	1092	-----SPLLPVSFLPTLIDLSLSQYSQFACELPRG-----APBLPOE-----G	1133
Db	871	CSPKHRNSSPRPR--HLAVADTKTLAPSE-----GEOPGIMPLPLASPLHLEVOGLIG	924
OY	1132	PTRRLSLEPGOLGALTSQ--PLHRHSGDPG	1158
Db	925	PCFS-SLEPHLGSVFKQDLDFQRHSGDPG	951

RESULT 15
ABP72167

ID	ABP72167	standard; Protein; 958	AA.
XX			
AC	ABP72167;		
XX			
DT	22-APR-2003	(first entry)	
XX			
DE	Human potassium channel 52908.		
XX			
KW	Human; potassium channel; 52908; nootropic; neuroprotective; analgesic; cardiovascular; cardiact; cyclostatic; immunosuppressive; antiinflammatory; osteopathic; hepatotropic; virocidic;		
KW	gene therapy.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Domain	41..59	
FT		/note= "PAS domain"	
FT	Domain	93..135	
FT		/note= "PAC domain"	
FT	Domain	261..282	
FT		/note= "Transmembrane domain"	
FT	Domain	295..317	
FT		/note= "Transmembrane domain"	
FT	Domain	345..361	
FT		/note= "Transmembrane domain"	
FT	Domain	371..390	
FT		/note= "Transmembrane domain"	
FT	Domain	398..420	
FT		/note= "Transmembrane domain"	
FT	Domain	491..515	
FT		/note= "Transmembrane domain"	
FT	Domain	341..580	
FT		/note= "Cyclic nucleotide gated channel transmembrane region"	
FT	Domain	463..482	
FT		/note= "Pore loop domain"	
FT	Region	478..480	
FT		/note= "potassium channel pore motif"	
FT	Region	608..699	
FT		/note= "Cyclic nucleotide gated channel transmembrane region"	
PN	MO2003006679-A2.		
XX			
PD	23-JAN-2003.		
XX			
PF	10-JUL-2002; 2002WO-US22456.		
XX			
PR	10-JUL-2001; 2001US-3042436.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Curtis RAD;		
XX			
DR	WPI; 2003-221759/21.		
XX			
DR	N-PSDB; ABZ58129.		
XX			
PT	New potassium channel family member, designated 52908, nucleic acids and proteins, useful for diagnosing, preventing, alleviating or treating e.g. pain and metabolic disorders, cardiovascular disorders or viral diseases	-	
PT			
XX			
PS	Claim 13; Page 100-102; 118pp; English.		
XX			
CC	The present sequence is the protein sequence of human 52908, a protein having the structural characteristics of a member of the potassium channel family. The invention provides 52908 nucleic acid molecules, recombinant expression vectors, host cells, transgenic animals in which a 52908 gene has been introduced or disrupted, 52908 proteins, fusion proteins, antigenic peptides and anti-52908 antibodies. The 52908 nucleic acids and proteins are useful in screening assays, predictive medicine (e.g. diagnostic		

CC and prognostic assays, monitoring clinical trials, and
 CC pharmacogenomics), and in methods of treatment of central nervous
 CC system (e.g. neurological disorders), pain and metabolic disorders,
 CC cardiovascular disorders, cellular proliferative, growth,
 CC differentiation and/or migration disorders, immune e.g.
 CC inflammatory disorders, disorders associated with bone metabolism,
 CC endothelial cell disorders, liver disorders, and viral diseases.

XX Sequence 958 AA:

Query Match 49.4%; Score 3001.5; DB 24; Length 958;

Best Local Similarity 53.7%; Pred. No. 3.8e-235; Mismatches 139; Indels 347; Gaps 30;

Matches 659; Conservative 83; Mismatches 139; Indels 347; Gaps 30;

QY 1 MPVRGHVAPONTFLDTIRKKEGSRKFIITANAVENCAYICNDGCELCGYSRAVYM 60
 DB 1 MPVRGHVAPONTFLDTIRKKEGSRKFIITANAVENCAYICNDGCELCGYSRAVYM 60
 QY 61 QRPCTDFLHGPRFORAAQIAQALLGAERKVEIAFYRKDGSCFLCLVDVVPYKNDG 120
 DB 61 QRPCTDFLHGPRFORAAQIAQALLGAERKVEIAFYRKDGSCFLCLVDVVPYKNDG 120
 QY 121 AVIMFIFNEFVMEKMDWGPADHTNHRGPTSMALPARAKTFRLKLALLA-LTARESS 179
 DB 121 AVIMFIFNEFVMEKMDWGPADHTNHRGPTSMALPARAKTFRLKLALLA-LTARESS 179
 QY 180 VR-----SGGAGGAGAPGAVVVDVLTLPAPSSSESLALDEVTAMDNHVGAPAEERRA 233
 DB 146 QRLLSQSFLLGSESHGRPG----- 164
 QY 234 LVGSGSPRSAPGQLPSPRAHSLNDASGSSCLARTSRSCASVRRASSADIEAMRA 293
 DB 165 --GPG-----PG----- 169
 QY 294 GVLPPRRHASTGAMHPLRSGLLNSTSDSLVRYRTISKIPQITLNFYDLKGDPELASPT 353
 DB 170 -----TGR-----GKRTTISQIPQITLNFVEFNLEKRRSSST 201
 QY 354 SDRELIAP-KIKERTHNTYKTVOLSLGADVLPEYKLOAPRIHRMTILHYSFKAVNDW 412
 DB 202 TEELIAPHKRYVERONTKEVTVLSGADVLPREYKLOAPRIHRMTILHYSFKAVNDW 261
 QY 413 LILLIYITAVFTPYSAFLKETEDEGPATECGYACOPLAVIDLIDIMFIVDILINFR 472
 DB 262 LILLIYITAVFTPYSAFLKETEDEGPATECGYACOPLAVIDLIDIMFIVDILINFR 320
 QY 473 TTYVANEEVYSHGRILAVHYFKGMELIDMVAALPFDLLIFRTGSDDETTLGLKTARL 529
 DB 321 TTYVANDEEVVSHPRILAVHYFKGMFLIDMVAALPFDLLIFRTGSDDETTLGLKTARL 380
 QY 530 LRLVRAKLDYRSEYGAANFLMCTPALLAHMLACIMVAIGNMEOPHMDSRIGMLNL 589
 DB 381 LRLVRAKLDYRSEYGAANFLMCTPALLAHMLACIMVAIGNMEOPHMDSRIGMLNL 440
 QY 590 GDQIGKPYNSG-LGGSIKDKYVYALYFTPSLTSVGFANVSPTNSEKIFPSICVMLIG 648
 DB 441 GVOIGKRYNSGDPASGSPVDKYYALYFTPSLTSVGFANVSPTNSEKIFPSICVMLIG 500
 QY 649 SLMTASTIGNSATIIRLYSGTARHTOMLRVERIFRHOJPNPLRORLEVEFOHAWSYT 708
 DB 501 SLMTASTIGNSATIIRLYSGTARHTOMLRVERIFRHOJPNPLRORLEVEFOHAWSYT 560
 QY 709 NGIDMNAVLFKGFPECLQADICLHLNRSLLQCHKPFGATKGCIRALAMKFTTHAPPGDT 768
 DB 561 NGIDMNAVLFKGFPECLQADICLHLNRSLLQCHKPFGATKGCIRALAMKFTTHAPPGDT 620
 QY 769 LVHAGDLTALYFISRSISIELRGDVVVALLGKNDIFGEPLNLVYARPGKSGDVVALTYC 828
 DB 621 LVHAGDLTALYFISRSISIELRGDVVVALLGKNDIFGEPLNLVYARPGKSGDVVALTYC 680
 QY 829 DLHKIHBDLLEVDIMYEPESDHFWSLEIFENLRDT-----NMTPGSPGSTELEGFSR 883
 DB 681 DLHKIHBDLLEVDIMYEPESDHFWSLEIFENLRDT-----NMTPGSPGSTELEGFSR 738

QY 884 QKRKLSFRRTKDXTEOPGEVSALGPBAGAPSSRGRPGGPMGESSPSGSPRESSED 943
 DB 739 -----LSDNGSD-----AAPPLISIDASGLM-----PELLOE 765
 QY 944 EGPGRSSPRLKLVPPSSRRPPGEPPELMEDECKSSDTCNPLSGAFSGVSNIFSWGD 1003
 DB 766 -----MPRHSPOSQEDP-----DC-----W-- 782
 QY 1004 SRGRQYELPCOPAPTPSLNIPLSSPGRPRGDVSRDLALQROLNLETRLSADMATV 1063
 DB 783 -----PLK-----LGSRLQOLQOAMNLESRYSSDLSRT 811
 QY 1064 LQLLQROM-----TLVPPAYSAVTPGP-----GPTSP----- 1091
 DB 812 LQLLQROMPGHASYTLEAPASNDLALVPIA-SETTSGPRLPGFLPPAQTPSYGDLDD 870
 QY 1092 -----SPLPSPPLTLTLDLSLOYQFMACELEPPG-----APELPOE-----G 1131
 DB 871 CSPKHNSSPRMP--HLAVATDKTLAPSS-----QOEPGGLMPLASPLHPLLEVQGLIG 924
 QY 1132 PTRRLSLPGOLGALTSQ-PLHRHGSDPG 1158
 DB 925 PCFS-SLPEHLGVSVPKQLDFORHGSDPG 951

Search completed: September 23, 2003, 17:37:01
 Job time : 81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 11:02:09 ; Search time 34 Seconds

(without alignments)
1603.057 Million cell updates/sec

Title: US-10-000-151b-3

Perfect score: 6079
Sequence: 1 MPVRRGHVAPONTFLDIIR.....GOLGALTQPLRHRSDDPGS 1159.

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6079	100.0	1159	1 KCH2_HUMAN	Q12809 homo sapien
2	5942	97.7	1161	1 KCH2_RABIT	Q8WNY2 oryctolagus
3	5872.5	96.6	1158	1 KCH2_CANFA	Q9TSZ3 canis famli
4	5797.5	95.4	1162	1 KCH2_MOUSE	Q35219 mus musculu
5	5790	95.2	1163	1 KCH2_RAT	Q08962 rattus norv
6	3312.5	54.5	1195	1 KCH7_RAT	Q54852 rattus norv
7	3283.5	54.0	1195	1 KCH7_MOUSE	Q9ER47 mus musculu
8	3255	53.5	1196	1 KCH7_HUMAN	Q9H250 homo sapien
9	3043.5	50.1	994	1 KCH6_HUMAN	Q9H252 homo sapien
10	2982	49.1	950	1 KCH6_RAT	Q54853 rattus norv
11	2275.5	37.4	526	1 KCH2_CHICK	Q9P184 gallus gall
12	1636	26.9	1017	1 KCH4_HUMAN	Q9UG05 homo sapien
13	1616	26.6	1017	1 KCH4_RAT	Q9R175 rattus norv
14	1568	25.8	1107	1 KCH8_HUMAN	Q96142 homo sapien
15	1551	25.5	1102	1 KCH8_RAT	Q9QWS8 rattus norv
16	1540.5	25.3	1083	1 KCH3_HUMAN	Q9ULI8 homo sapien
17	1537.5	25.3	1087	1 KCH3_RAT	Q89047 rattus norv
18	1527.5	25.1	1087	1 KCH3_MOUSE	Q9WVJ0 mus musculu
19	1448.5	23.8	1174	1 CIKE_DROME	Q02280 drosophila
20	1409.5	23.2	962	1 KCH1_RAT	Q63472 rattus norv
21	1406.5	23.1	987	1 KCH1_BOVIN	Q18965 bos taurus
22	1402.5	23.1	989	1 KCH1_HUMAN	Q95259 homo sapien
23	1402	23.1	989	1 KCH1_MOUSE	Q60603 mus musculu
24	1399.5	23.0	988	1 KCH5_HUMAN	Q8HMC2 homo sapien
25	1398.5	23.0	988	1 KCH5_RAT	Q9EP19 rattus norv
26	1193	19.6	876	1 KCH8_MOUSE	P99111 mus musculu
27	804	13.2	162	1 KCH4_CAVPO	Q8V66 oryctolagus
28	595.5	9.8	1175	1 KCH4_RABIT	Q9TKAT rattus norv
29	581	9.6	1198	1 KCH4_MOUSE	Q70507 mus musculu
30	578	9.5	1186	1 KCH4_HUMAN	Q9Y344 homo sapien
31	563	9.3	1203	1 KCH2_HUMAN	Q9U151 homo sapien
32	551.5	9.1	889	1 HCN2_MOUSE	Q88703 mus musculu
33	543	8.9	863	1 HCN2_MOUSE	Q88703 mus musculu

34	529.5	8.7	822	1 HCN1_RABIT	Q8WZS1 oryctolagus
35	527.5	8.7	890	1 HCN1_HUMAN	Q60741 homo sapien
36	524.5	8.6	834	1 HCN2_RAT	Q9TKA9 rattus norv
37	515.5	8.5	910	1 HCN1_MOUSE	Q88704 mus musculu
38	512.5	8.4	774	1 HCN3_HUMAN	Q9P123 homo sapien
39	509.5	8.4	780	1 HCN3_RAT	Q9TKA8 rattus norv
40	507.5	8.3	910	1 HCN1_RAT	Q9TKB0 rattus norv
41	506.5	8.3	779	1 HCN3_MOUSE	Q88705 mus musculu
42	476	7.8	97	1 KCH2_PIG	Q9TUI4 sus scrofa
43	451.5	7.4	665	1 CNG3_DROME	Q24278 drosophila
44	440	7.2	706	1 CNG3_BOVIN	Q29441 bos taurus
45	438	7.2	694	1 CNG3_HUMAN	Q16281 homo sapien

ALIGNMENTS

RESULT 1
KCH2_HUMAN STANDARD; PRT; 1159 AA.
ID Q12809; 075618; 075680; Q9HFT7; Q9HUT7; Q9H3P0.
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go
DE related gene potassium channel 1) (H-ERG) (ERG1) (Ether-a-go-go
DE related protein 1) (Eag related protein 1) (eag homolog).
GN KCHH2 OR HERG OR HERG1 OR ERG OR ERG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
[1] SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Hippocampus;
RX MEDLINE=94211879; PubMed=8159766;
RA Warneke J.W., Ganetzky B.;
RT "A family of potassium channel genes related to eag in Drosophila and
RT mammals".
RL Proc. Natl. Acad. Sci. U.S.A. 91:3438-3442(1994).
[2]
[2] SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT LGT2 CYS-534.
RX MEDLINE=98260867; PubMed=9600240;
RA Itoh T., Tanaka T., Nagai R., Kamita T., Sawayama T., Nakayama T.,
RA Tomokita H., Sakurada H., Yazaki Y., Nakamura Y.;
RT "Genomic organization and mutational analysis of HERG, a gene
RT responsible for familial long QT syndrome".
RL Hum. Genet. 102:435-439(1998).
[3]
[3] SEQUENCE FROM N.A. (ISOFORM 1).
RA Downie D., Chapman C.G., Pinta P., Rice S., Bahmani F., Mirdock P.,
RA Pearson N., Randall A.D., Meadows H.J.;
RT "Potent inhibition of HERG K⁺ channels by the neuroprotective agent
RT Sildenafil".
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[4]
[4] SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Heart;
RX MEDLINE=21269186; PubMed=11374908;
RA Soejima H., Kawamoto S., Akai J., Miyoshi O., Arai Y., Morohka T.,
RA Matsuo S., Nakawa N., Kimura A., Okubo K., Mukai T.;
RT "Isolation of novel heart-specific genes using BodyMap database".
RL Genomics 74:115-120(2001).
[5]
[5] SEQUENCE FROM N.A. (ISOFORM 2).
RA Crociani O., Guasti L., Balzi M., Becchetti A., Wanke E., Olivetto M.,
RA Wyome R.S., Arcangeli A.;
RT "Cell cycle-dependent expression of HERG1 and HERG1b isoforms in tumor
RT cells".
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
[6]
[6] SEQUENCE OF 1-376 FROM N.A. (ISOFORM 2).
RC TISSUE=Heart;

RA MEDLINE-98012815; PubMed-9351462;
 RA London B., Trudeau M.C., Newton K.P., Beyer A.K., Copeland N.G.,
 RA Gilbert D.J., Jenkun N.A., Satler C.A., Robertson G.A.;
 RT "Two isoforms of the mouse ether-a-go-go-related gene coassemble to
 RT form channels with properties similar to the rapidly activating
 RT component of the cardiac delayed rectifier K⁺ current.";
 RL Circ. Res. 81:870-878(1997).
 RN [17]
 RN SEQUENCE OF 1-376 FROM N.A. (ISOFORM 2).
 RC TISSUE=Heart atrium;
 RX MEDLINE-98012799; PubMed-9351446;
 RA Lees-Miller J.P., Kondo C., Wang L., Duff H.J.;
 RT "Electrophysiological characterization of an alternatively processed
 RT ERG K⁺ channel in mouse and human hearts.";
 RL Circ. Res. 81:719-726(1997).
 RN [18]
 RN SEQUENCE OF 27-1159 FROM N.A. (ISOFORM 1).
 RA Yang P., Paulussen A., Verhasselt P., Crabbe R., Luyten W.,
 RA Armstrong M.;
 RT "Analysis of the human HERG gene: Intron localisation and
 RT identification of a novel inherited mutation associated with long
 RT QT.";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [19]
 RN SEQUENCE OF 795-1159 FROM N.A. (ISOFORM 3).
 RC TISSUE=Heart ventricle;
 RX MEDLINE-98438490; PubMed-9765245;
 RA Kupersmidt S., Snyder D.J., Raes A., Roden D.M.;
 RT "A K⁺ channel splice variant common in human heart lacks a C-terminal
 RT domain required for expression of rapidly activating delayed
 RT rectifier current.";
 RL J. Biol. Chem. 273:27231-27235(1998).
 RN [110]
 RN SEQUENCE OF 59-1159 FROM N.A. (ISOFORM 4), AND SEQUENCE OF 711-1159
 RP FROM N.A. (ISOFORM 1/2).
 RC TISSUE=Brain;
 RX MEDLINE-22388257; PubMed-12477932;
 RA Strassberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stadelman M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosok S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [111]
 RN MUTAGENESIS OF ASN-598; ASN-629 AND SER-631, AND N-GLYCOSYLATION.
 RX MEDLINE-22057172; PubMed-12063277;
 RA Gong Q., Anderson C.L., January C.T., Zhou Z.;
 RT "Role of glycosylation in cell surface expression and stability of
 RT HERG potassium channels.";
 RL Am. J. Physiol. 283:H77-H84(2002).
 RN [112]
 RN MUTAGENESIS OF SER-283; SER-890; THR-895 AND SER-1137, AND
 RP PHOSPHORYLATION.
 RX MEDLINE-20299343; PubMed-10837251;
 RA Cui J., Melman Y., Palma E., Fishman G.I., McDonald T.V.;
 RT "Cyclic AMP regulates the HERG K(+) channel by dual pathways.";
 RL Curr. Biol. 10:671-674(2000).
 RN [113]
 RN INTERACTION WITH KCNE1.

RX MEDLINE-97373956; PubMed-9230439;
 RA McDonald T.V., Yu Z., Ming Z., Palma E., Meyers M.B., Wang K.-W.,
 RA Goldstein S.A., Fishman G.I.;
 RT "A link-HERG complex regulates the cardiac potassium current I(Kr).";
 RL Nature 388:289-292(1997).
 RN [114]
 RN INTERACTION WITH KCNE2.
 RX MEDLINE-99235979; PubMed-10219239;
 RA Abbott G.W., Sesti F., Splawski I., Buck M.E., Lehmann M.H.,
 RA Timothy K.W., Keating M.T., Goldstein S.A.;
 RT "MiRP1 forms IKr potassium channels with HERG and is associated with
 RT cardiac arrhythmia";
 RL Cell 97:175-187(1999).
 RN [115]
 RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-135, AND MUTAGENESIS OF
 RP PHE-29 AND TYR-43.
 RX MEDLINE-99059500; PubMed-9845367;
 RA Morris Cabral J.H., Lee A., Cohen S.L., Chait B.T., Li M.,
 RA Mackinnon R.;
 RT "Crystal structure and functional analysis of the HERG potassium
 RT channel N terminus: a eukaryotic PAS domain.";
 RL Cell 95:649-655(1998).
 RN [116]
 RN VARIANTS LOT2 ASP-470; VAL-561 AND SER-628.
 RX MEDLINE-95196272; PubMed-7889573;
 RA Curran M.E., Splawski I., Timothy K.W., Vincent G.M., Green E.D.,
 RA Keating M.T.;
 RT "A molecular basis for cardiac arrhythmia: HERG mutations cause long
 RT QT syndrome.";
 RL Cell 80:795-803(1995).
 RN [117]
 RN VARIANT LOT2 MET-822.
 RX MEDLINE-97071892; PubMed-8914737;
 RA Satler C.A., Walsh E.P., Vesely M.R., Plummer M.H., Ginsburg G.S.,
 RA Jacob H.J.;
 RT "Novel missense mutation in the cyclic nucleotide-binding domain of
 RT HERG causes long QT syndrome.";
 RL Am. J. Med. Genet. 65:27-35(1996).
 RN [118]
 RN VARIANT LOT2 ARG-593.
 RX MEDLINE-96259954; PubMed-8635257;
 RA Benson D.W., Macrae C.A., Vesely M.R., Walsh E.P., Seidman J.G.,
 RA Seidman C.E., Satler C.A.;
 RT "Missense mutation in the pore region of HERG causes familial long QT
 RT syndrome.";
 RL Circulation 93:1791-1795(1996).
 RN [119]
 RN VARIANT LOT2 THR-561.
 RX MEDLINE-97031865; PubMed-8877771;
 RA Dausse E., Berthet M., Denjoy I., Andre-Fouet X., Crouad C.,
 RA Benacœur M., Faure S., Coumel P., Schwartz K., Guicheney P.;
 RT "A mutation in HERG associated with notched T waves in long QT
 RT syndrome.";
 RL J. Mol. Cell. Cardiol. 28:1609-1615(1996).
 RN [120]
 RN VARIANTS LOT2 ILE-474; HIS-611; VAL-614 AND LEU-630.
 RX MEDLINE-97176600; PubMed-9024139;
 RA Tanaka T., Nagai R., Tomoike H., Takata S., Yano K., Yabuta K.,
 RA Haneida N., Nakano O., Shibata A., Sawayama T., Kasai H., Yazaki Y.,
 RA Nakamura Y.;
 RT "Four novel KVLQT1 and four novel HERG mutations in familial long-QT
 RT syndrome.";
 RL Circulation 95:565-567(1997).
 RN [121]
 RN VARIANTS LOT2 CYS-572; ASP-588; VAL-614 AND ALA-630.
 RX MEDLINE-98360095; PubMed-9693036;
 RA Splawski I., Shen J., Timothy K.W., Vincent G.M., Lehmann M.H.,
 RA Keating M.T.;
 RT "Genomic structure of three long QT syndrome genes: KVLQT1, HERG, and
 RT KCNE1.";
 RL Genomics 51:86-97(1998).
 RN [122]
 RN VARIANTS LOT2 LEU-612; VAL-614; ASP-629; SER-629 AND SER-633.

Query Match 100.0%; Score 6079; DB 1; Length 1159;
 Best Local Similarity 100.0%; Pred. No. 3.3e-307;
 Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MPVRGHVAPQNTFLDTIIRKEFGQSRKFTIANARVENCAYIYCNDFCELCGYSRAEVM 60
DB 1 MPVRGHVAPQNTFLDTIIRKEFGQSRKFTIANARVENCAYIYCNDFCELCGYSRAEVM 60
QY 61 QRCCTDDELHGPRTORRAAQIAOALLGAERKEVEIAFYKDDSCFCLVDVVPVKNEDG 120
DB 61 QRCCTDDELHGPRTORRAAQIAOALLGAERKEVEIAFYKDDSCFCLVDVVPVKNEDG 120
QY 121 AVIMFLINFEVWEKMDVGSPPADTNHRCPTSMALPGRKATPRKLPALLATLARESSV 180
DB 121 AVIMFLINFEVWEKMDVGSPPADTNHRCPTSMALPGRKATPRKLPALLATLARESSV 180
QY 181 RSGGAGAGAGAVVVDVLTTPAASSESLALDEVTAMDNVAGLGPABERRALVGRSP 240
DB 181 RSGGAGAGAGAVVVDVLTTPAASSESLALDEVTAMDNVAGLGPABERRALVGRSP 240
QY 241 PRAPQOLPSPRAHSLNPDASGSSCLARTRESGASVYRASSADDIEMRAGVLP 300
DB 241 PRAPQOLPSPRAHSLNPDASGSSCLARTRESGASVYRASSADDIEMRAGVLP 300
QY 301 RHASTGAMHPLRGLNSTSDSLVRYRTISKIPQTLNFVDLKGDFLSPSDREITIA 360
DB 301 RHASTGAMHPLRGLNSTSDSLVRYRTISKIPQTLNFVDLKGDFLSPSDREITIA 360
QY 361 PKIKERTHNTKVTQVLSGADVLDEYKLOAPRIHMTILHYSPEKAVMDLILLVIY 420
DB 361 PKIKERTHNTKVTQVLSGADVLDEYKLOAPRIHMTILHYSPEKAVMDLILLVIY 420
QY 421 TAVFTYSAFLKETEEREGPATECGACOPLAVDLIVIMFLVDLINFRTTYVANE 480
DB 421 TAVFTYSAFLKETEEREGPATECGACOPLAVDLIVIMFLVDLINFRTTYVANE 480
QY 481 EYVSHFGRIAVHYFKGWFLLDMYAAIPEDLLIFSGSEELIGLKTARLLRVAVARKLD 540
DB 481 EYVSHFGRIAVHYFKGWFLLDMYAAIPEDLLIFSGSEELIGLKTARLLRVAVARKLD 540
QY 541 RYSEYGAANVELLMCTFPALIAHMLACTIWAIGNMOPHMDSRIGMLHNLDOIGKPYNS 600
DB 541 RYSEYGAANVELLMCTFPALIAHMLACTIWAIGNMOPHMDSRIGMLHNLDOIGKPYNS 600
QY 601 GIGGPRSTKDKYVVALYFTFSSLSVSGVNSPMTNSKIFSTCYMLIGSLMYASIFGNVS 660
DB 601 GIGGPRSTKDKYVVALYFTFSSLSVSGVNSPMTNSKIFSTCYMLIGSLMYASIFGNVS 660
QY 661 AIIORLUSGARVHTOMLRVREFIRFHOIPNPLRORLEEFQOHAWSYTGIDMNAVYKGF 720
DB 661 AIIORLUSGARVHTOMLRVREFIRFHOIPNPLRORLEEFQOHAWSYTGIDMNAVYKGF 720
QY 721 PECLQADICLHNLKRSLLQHCPRGATKGLRALAMKFTTHAPGPDIVLAGDULITALY 780
DB 721 PECLQADICLHNLKRSLLQHCPRGATKGLRALAMKFTTHAPGPDIVLAGDULITALY 780
QY 781 FIERGSTIELRGVVAAILLKNIDIFGEPMLYARPGSNDDVALYTCDLHKIHRDLE 840
DB 781 FIERGSTIELRGVVAAILLKNIDIFGEPMLYARPGSNDDVALYTCDLHKIHRDLE 840
QY 841 VLMYPEFSDHFWSSLETFNLDNTNIPGSPSTELGEGFSQORRKLSEFRRTDQTE 900
DB 841 VLMYPEFSDHFWSSLETFNLDNTNIPGSPSTELGEGFSQORRKLSEFRRTDQTE 900
QY 901 QPGEVSLGPGRAGAGPSRGRPGGPGWGESPPSGPSSSEDEGPGRSSPRLVFPSS 960
DB 901 QPGEVSLGPGRAGAGPSRGRPGGPGWGESPPSGPSSSEDEGPGRSSPRLVFPSS 960
QY 961 PRPPGPPGPPPLMEDECEKSSDTCNPLSGAFSGVSNIFSWGDSRGQYDELPRCAPTP 1020
DB 961 PRPPGPPGPPPLMEDECEKSSDTCNPLSGAFSGVSNIFSWGDSRGQYDELPRCAPTP 1020

```

```

QY 1021 SLNINPLSSGRRRGVDSERDALQRLNRLRLTSADNAVYLQLLQRMQLVPPAYSA 1080
DB 1021 SLNINPLSSGRRRGVDSERDALQRLNRLRLTSADNAVYLQLLQRMQLVPPAYSA 1080
QY 1081 VTTTPGPPPTSTPLPSPVPTLTLDLSQVSGFMACEELPPGAPELPQSGPPTRLSLPG 1140
DB 1081 VTTTPGPPPTSTPLPSPVPTLTLDLSQVSGFMACEELPPGAPELPQSGPPTRLSLPG 1140
QY 1141 QLGALTSQPLHRHSDPGS 1159
DB 1141 QLGALTSQPLHRHSDPGS 1159

RESULT 2
KCH2_RABIT
ID KCH2_RABIT STANDARD; PRT: 1161 AA.
AC Q8WN12; 002731; 019119; 097586; 09Y06;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go
DE related gene potassium channel 1) (ERG1) (RERG) (ra-ery) (ether-a-go-
DE go related protein 1) (Eag related protein 1).
GN KCNH2 OR ERG.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA Wlitchel H.J., Hancock J.C., Levi A.J., Meech R.M.;
RT "RERG - rabbit ventricular ERG K+ channel subunit."
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 411-571 FROM N.A. (ISOFORM 1/2).
RX MEDLINE=97164986; Pubmed=9012748;
RA Wymore R.S., Gintant G.A., Wymore R.T., Dixon J.E., McKinnon D.,
RA Cohen I.S.;
RT "Issue and species distribution of mRNA for the IKr-like K+ channel,
RT erg."
RL Circ. Res. 80:261-268(1997).
CC CC
CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly
CC rectifying potassium channel. Channel properties are modulated by
CC cAMP and subunit assembly. Mediates the rapidly activating
CC component of the delayed rectifying potassium current in heart
CC (IKr) (by similarity).
CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
CC heterotetrameric complex of pore-forming alpha subunits that can
CC associate with modulating beta subunits. Heteromultimer with
CC KCNH6/ERG2, KCNH7/ERG3, KCNE1 and KCNE2 (by similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms-2:
CC Name=1;
CC IsoId=Q8WNV2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8WNV2-2; Sequence=VSP_000971;
CC CC
CC -1- TISSUE SPECIFICITY: Detected in heart, both in atrium and in left
CC ventricle.
CC -1- DOMAIN: The segment 54 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -1- PTM: Phosphorylated on serine and threonine residues (by
CC similarity).
CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)
CC subfamily.
CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U87513; AAB68612.1; -
 DR EMBL: AF068736; AAC99425.1; -
 DR EMBL: AF05061; AAD39357.1; -
 DR EMBL: U75212; AAC48723.1; -
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nig.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding.1.
 DR Pfam: PF00520; Ion_trans.1.
 DR Pfam: PF00785; PAC.1.
 DR SMART: SM00100; CNMP.1.
 DR SMART: SM00086; PAC.1.
 DR PROSITE: PS50042; CNMP_BINDING_3; 1.
 DR PROSITE: PS50112; PAS.1.
 DR PROSITE: PS50113; PAC.1.
 KM Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KM Potassium channel; Potassium; Potassium transport; Transmembrane;
 KM Phosphorylation; Glycoprotein; Multigene family; Alternative splicing.
 FT DOMAIN 1 405
 FT TRANSMEM 406 426
 FT TRANSMEM 453 473
 FT DOMAIN 474 497
 FT TRANSMEM 498 518
 FT TRANSMEM 523 543
 FT DOMAIN 544 549
 FT TRANSMEM 550 570
 FT DOMAIN 614 634
 FT TRANSMEM 641 661
 FT DOMAIN 662 1161
 FT DOMAIN 17 88
 FT DOMAIN 92 144
 FT DOMAIN 299 302
 FT NP_BIND 744 861
 FT CARBOHYD 600 600
 FT VANSPLIC 69 85
 FT CONFLICT 411 411
 FT CONFLICT 445 446
 FT CONFLICT 553 553
 FT CONFLICT 561 561
 FT SEQUENCE 1161 AA; 126961 MW; 798532B2FFBD9AEB CRC64;

Query Match 97.7%; Score 5942; DB 1; Length 1161;
 Best Local Similarity 98.0%; Pred. No. 4.1e-300;
 Matches 1138; Conservative 2; Mismatches 19; Indels 2; Gaps 1;

QY 1 MPVRGHVAPONTFLDTIIRKFEQOSRKFIITANAENVCAVYICNDGFCGLCGYSRAEVM 60
 DB 1 MPVRGHVAPONTFLDTIIRKFEQOSRKFIITANAENVCAVYICNDGFCGLCGYSRAEVM 60

QY 61 ORPCTCDLHGPRTORRAAIOAALLGAERKVEIATYRKDGSFCLLVVVPKKNDDG 120
 DB 61 ORPCTCDLHGPRTORRAAIOAALLGAERKVEIATYRKDGSFCLLVVVPKKNDDG 120

QY 121 AVIMEILNFEVMEKDMVSPAHDTNHRGPTSMILAPRAKTFRLKLALLATARSSV 180
 DB 121 AVIMEILNFEVMEKDMVSPAHDTNHRGPTSMILAPRAKTFRLKLALLATARSSV 180

QY 121 AVIMEILNFEVMEKDMVSPAHDTNHRGPTSMILAPRAKTFRLKLALLATARSSV 180
 DB 121 AVIMEILNFEVMEKDMVSPAHDTNHRGPTSMILAPRAKTFRLKLALLATARSSV 180

QY 181 RSGAGAGAGAGAVVVDLTPAAPSSSESLADEVTANDNHNVAIGLPAEERRALVPGSP 240
 DB 181 RSGAGAGAGAGAVVVDLTPAAPSSSESLADEVTANDNHNVAIGLPAEERRALVPGSP 240

QY 241 PR--SAPGQLPSPRAHSLNPDASSGSCSLANTRSCASVRRASADDIEMRAGVLP 298
 DB 241 PR--SAPGQLPSPRAHSLNPDASSGSCSLANTRSCASVRRASADDIEMRAGVLP 298

DB 241 PR--SAPGQLPSPRAHSLNPDASSGSCSLANTRSCASVRRASADDIEMRAGVLP 298
 QY 299 PPRHASTGAMHPLRGLNSTSDSLVRYRTSKIPQITLNEVFDKGPFLASPSDBEI 358
 DB 301 PPRHASTGAMHPLRGLNSTSDSLVRYRTSKIPQITLNEVFDKGPFLASPSDBEI 360

QY 359 IAPKIKERTHNHYTEKTOVLSGADVLEPYKQAPRIHRMTILHSPKAVWMDLILLY 418
 DB 361 IAPKIKERTHNHYTEKTOVLSGADVLEPYKQAPRIHRMTILHSPKAVWMDLILLY 420

QY 419 IYTAFTPYSAFLKETEEDGPAECGYACQPLAVDLYDIMIYDILINFRTYVNA 478
 DB 421 IYTAFTPYSAFLKETEEDGPAECGYACQPLAVDLYDIMIYDILINFRTYVNA 480

QY 479 NEEVSHHGRLAVHFKGMFLDMVAALPFDLLIGSGSEELIGLKTARLLRLTVARK 538
 DB 481 NEEVSHHGRLAVHFKGMFLDMVAALPFDLLIGSGSEELIGLKTARLLRLTVARK 540

QY 539 LDRSEYGAAYFLIMCTFALIAHMLACIWAIGMEOPHMDSRIGMLHNGDQIKDPYN 598
 DB 541 LDRSEYGAAYFLIMCTFALIAHMLACIWAIGMEOPHMDSRIGMLHNGDQIKDPYN 600

QY 599 SSGLGSPGSIKQYVATLFTSSSLTVSGRANVSPNTSEKIFSIQVMILGSLMVAISTGN 658
 DB 601 SSGLGSPGSIKQYVATLFTSSSLTVSGRANVSPNTSEKIFSIQVMILGSLMVAISTGN 660

QY 659 VSAIIQRLYSGTARYHTOMLAVREFIRFHQIIPNPLRQRLSEYFOHAMSYTNGIDNNAYLK 718
 DB 661 VSAIIQRLYSGTARYHTOMLAVREFIRFHQIIPNPLRQRLSEYFOHAMSYTNGIDNNAYLK 720

QY 719 GFPECLQADICLHLNRSILOHCKPRGATKGCRLALAMKFKTTNAPPGDYLVAHAGDLLTA 778
 DB 721 GFPECLQADICLHLNRSILOHCKPRGATKGCRLALAMKFKTTNAPPGDYLVAHAGDLLTA 780

QY 779 LYFISRGSEILIRGTVVAAIIGKNDIGCEPLNTVAPRCKSGNDVVALYYCDLHKIHRDIL 838
 DB 781 LYFISRGSEILIRGTVVAAIIGKNDIGCEPLNTVAPRCKSGNDVVALYYCDLHKIHRDIL 840

QY 839 LEVLDMPYEFSDHEWSSLEITFNLRDTNMIPGSPSTLEGGFSRQRKRLSFRRTDND 898
 DB 841 LEVLDMPYEFSDHEWSSLEITFNLRDTNMIPGSPSTLEGGFSRQRKRLSFRRTDND 900

QY 899 TEQGEVSALGPRGAGPSSRGPRGPGWGESPPSSGESSDEDEGGRSSPLRLVPF 958
 DB 901 TEQGEVSALGPRGAGPSSRGPRGPGWGESPPSSGESSDEDEGGRSSPLRLVPF 960

QY 959 SSPRRGPRGGEPLMECEKSSDTCNPLSGAFSGVSNIFSFWDKSRGROYELPRCAP 1018
 DB 961 SSPRRGPRGGEPLMECEKSSDTCNPLSGAFSGVSNIFSFWDKSRGROYELPRCAP 1020

QY 1019 TPSSLINIPSSGRRPRGSDVESRLDALQORNLRETRLNSADMAVYIQLQORMTLVPPAY 1078
 DB 1021 TPSSLINIPSSGRRPRGSDVESRLDALQORNLRETRLNSADMAVYIQLQORMTLVPPAY 1080

QY 1079 SAVTTPEGFPSTSPFLVSPPLTLDLSQVSOQFMAACEBLPGAPLPOEGPTRRLSL 1138
 DB 1081 SAVTTPEGFPSTSPFLVSPPLTLDLSQVSOQFMAACEBLPGAPLPOEGPTRRLSL 1140

QY 1139 PGOLGALTSOPLHRRGSDPS 1159
 DB 1141 PGOLGALTSOPLHRRGSDPS 1161

RESULT 3
 KCH2_CANFA STANDARD; PRT; 1158 AA.
 AC Q9T5Z3; 002719; 018820;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go
 related gene potassium channel 1) (ERGI) (c-erg) (DERG) (Ether-a-go-

QY 541 RYSEYGAAYFLMCTFALLAHMLACIWAYAIGNMEOPHMDSRIGWLNHNGOIGKPNSS 600
 DB 540 RYSEYGAAYFLMCTFALLAHMLACIWAYAIGNMEOPHMDSRIGWLNHNGOIGKPNSS 599
 QY 601 GLGSPSTKDKYVTLALYFFSSLSVGRGNVSPNTNSEKIFSCWMLGSLMYASTFGVNS 660
 DB 600 GLGSPSTKDKYVTLALYFFSSLSVGRGNVSPNTNSEKIFSCWMLGSLMYASTFGVNS 659
 QY 661 AIIORLXSGTARYTOMLRVREFIRFHQIPNPLRQRLSEYEQHMASTYNGIDMANVLKGF 720
 DB 660 AIIORLXSGTARYTOMLRVREFIRFHQIPNPLRQRLSEYEQHMASTYNGIDMANVLKGF 719
 QY 721 PECIADICHLNLSLQHCXKPPRGATKCLRALAMKFKTTHAPGGTIVHAGDILLALY 780
 DB 720 PECIADICHLNLSLQHCXKPPRGATKCLRALAMKFKTTHAPGGTIVHAGDILLALY 779
 QY 781 FISGSTEILRGDVVVALGKNDIFGEPLNIYARPGKNGVVALTYCDLKHRRDLE 840
 DB 780 FISGSTEILRGDVVVALGKNDIFGEPLNIYARPGKNGVVALTYCDLKHRRDLE 839
 QY 841 VLDVYPERSDHFMSSLETFENLRDTNMI PGSPGSTELEGGFSRORRKRKLSFRRTDKDTE 900
 DB 840 VLDVYPERSDHFMSSLETFENLRDTNMI PGSPGSALEGGFNROKKRKLSTRRTDRDPE 899
 QY 901 QPGEVSALGPGRAGAPSSSRGPRGPGWGESPPSSGPPSESDSDGPGRSSSPLRLVPRSS 960
 DB 900 QPGEVSALGPGRAGAPSSSRGPRGPGWGESPPSSGPPSESDSDGPGRSSSPLRLVPRSS 959
 QY 961 PRPGEPPGPGPRLMEDCKSSDTCNPISGASGVSNIFSPMGDSRGROYOLPRCPAPTP 1020
 DB 960 PRPGEPPGPGPRLMEDCKSSDTCNPISGASGVSNIFSPMGDSRGROYOLPRCPAPTP 1019
 QY 1021 SLNPIPLSPRRPRRGDVEESRLDALQORNLRLSADMATVYQLORQTVLPAYSA 1080
 DB 1020 SLNPIPLSPRRPRRGDVEESRLDALQORNLRLSADMATVYQLORQTVLPAYSA 1079
 QY 1081 VTPPGPPTSTPLPVSPPLTLTLDLSISOVSQPMACEELPPGAPELPQEGPTRLPLPG 1140
 DB 1080 VTPPGPPTSTPLPVSPPLTLTLDLSISOVSQPMACEELPPGAPELPQEGPTRLPLPG 1139
 QY 1141 QLGLATISOPILRHGSDPGS 1159
 DB 1140 QLGLATISOPILRHGSDPGS 1158

RESULT 4

KCH2_MOUSE STANDARD; PRT: 1162 AA.
 AC 035219; 035220; 035221; 035989;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-FEB-2003 (Rel. 42, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go related gene potassium channel 1) (ERGL1) (MERG1) (Ether-a-go-go related protein 1) (Eag related protein 1).
 GN KCH2 OR ERG OR MERG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBT_TaxID=10090;
 RX MEDLINE=98012815; PubMed=9351462;
 RA London B., Trudeau M.C., Newton K.P., Beyer A.K., Copeland N.G., Gilbert D.J., Jenkins N.A., Sailer C.A., Robertson G.A.;
 RT "Two isoforms of the mouse ether-a-go-go-related gene coassemble to form channels with properties similar to the rapidly activating RT component of the cardiac delayed rectifier K⁺ current.";
 CL Circ. Res. 81:870-878(1997).
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RC TISSUE=atrial tumor;
 RX MEDLINE=98012799; PubMed=9351446;
 RA Lees-Miller J.P., Rondo C., Wang L., Duff H.J.;
 RT "Electrophysiological characterization of an alternatively processed ERG K⁺ channel in mouse and human hearts.";
 CL Circ. Res. 81:719-726(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-186.
 RC TISSUE=Colon smooth muscle;
 RA Shoen F., Malykhina A., Akbarali H.I.;
 RT "Smooth muscle KCNH2 erg potassium channel.";
 RL Submitted (Jul-2002) to the EMBL/Genbank/DBJ databases.
 CC -i FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly rectifying potassium channel. Channel properties are modulated by cAMP and subunit assembly. Mediates the rapidly activating component of the delayed rectifying potassium current in heart (IKr) (By similarity).
 CC -i SUBUNIT: The potassium channel is probably composed of a homo- or heterotetrameric complex of pore-forming alpha subunits that can associate with modulating beta subunits. Heteromultimer with KCNH6/ERG2, KCNH7/ERG3, KCNE1 and KCNE2 (By similarity).
 CC -i SUBCELLULAR LOCATION: Integral membrane protein.
 CC -i ALTERNATIVE PRODUCTS: Integral membrane protein.
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Experimental confirmation may be lacking for some isoforms;
 CC Name=1; Synonyms=1A, A;
 CC IsoId=035219-1; Sequence=Displayed;
 CC Name=2; Synonyms=1A';
 CC IsoId=035219-2; Sequence=VSP_000969;
 CC Name=3; Synonyms=1B, B;
 CC IsoId=035219-3; Sequence=VSP_000970;
 CC -i TISSUE SPECIFICITY: Isoform 1 is expressed in heart, brain and testis and at low levels in lung. Isoform 3 is expressed predominantly in heart. The expression of isoform 2 is low in all tissues tested.
 CC -i DOMAIN: The segment S4 is probably the voltage-sensor and is characterized by a series of positively charged amino acids at every third position.
 CC -i PTM: Phosphorylated on serine and threonine residues (By similarity).
 CC -i SIMILARITY: Belongs to the potassium channel family. H (Eag) subfamily.
 CC -i SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -i SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -i SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -i CAUTION: Ref.3 sequence was originally reported as deriving from rabbit.
 CC -i CAUTION: Ref.3 sequence differs from that shown due to a frameshift in position 1057.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF012868; AAC3418.1; -;
 DR EMBL; AF012869; AAC3419.1; -;
 DR EMBL; AF012871; AAC3420.1; -;
 DR EMBL; AF012870; AAC3420.1; JOINED.
 DR EMBL; AF012871; AAC3421.1; -;
 DR EMBL; AF012871; AAC3422.1; -;
 DR EMBL; AF012870; AAC3422.1; JOINED.
 DR EMBL; AF034762; AAB87571.1; -;
 DR EMBL; AF439342; AAL3537.2; ALT_FRAME.
 DR MGD; MGI:1341722; Kcnh2.
 DR InterPro; IPR000595; CNMP_binding.
 DR InterPro; IPR003967; Erg_channel.
 DR InterPro; IPR005821; Ion_trans.

DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding_1.
 DR Pfam: PF00520; ion_trans_1.
 DR Pfam: PF00785; PAC_1.
 DR SMART: SM00100; CNMP_1.
 DR SMART: SM00086; PAC_1.
 DR PROSITE: PS50042; CNMP_BINDING_3_1.
 DR PROSITE: PS50112; PAC_1.
 DR PROSITE: PS50113; PAC_1.
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Phosphorylation; Glycoprotein; Multigene family; Polymorphism;
 KW Alternative splicing.
 FT DOMAIN 1 405
 FT TRANSMEM 406 426 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 453 473 SEGMENT S1 (POTENTIAL).
 FT DOMAIN 474 497 SEGMENT S2 (POTENTIAL).
 FT TRANSMEM 498 518 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 523 543 SEGMENT S3 (POTENTIAL).
 FT TRANSMEM 544 549 SEGMENT S4 (POTENTIAL).
 FT TRANSMEM 550 570 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 614 634 SEGMENT S5 (POTENTIAL).
 FT TRANSMEM 641 661 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 662 1162 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 17 88 PAS.
 FT DOMAIN 92 144 PAC.
 FT DOMAIN 299 302 POLY-PRO.
 FT NP_BIND 744 861 CNMP.
 FT CARBOHYD 600 600 Missing (in isoform 2).
 FT VARSPIC 1 59 /ftid=VSP_000969.
 FT VARSPIC 1 378 /ftid=VSP_000969.
 FT MPVRGHAAPONTFLDTIIRKFEQSGRRFTIANARVENCAY
 IYCDGFCCELCYSRAEMORPCYDHLGHPRTORRAAOI
 AOAALGAERKVEIAFYRKDSCFLCLVDVAVYVNEGAYI
 MFLTFEVMEKMGVSPAHDTNHRGPTSMLAGRAKTR
 LKIPALLALTARRESSVTGSMHSGAGAVYVDYDLPAAP
 SSESALDEVAMDNHVAAGLPAERATLVPGSASVAST
 RGPPSPRAQSLNDAGSSCSLARTSRSCASVRASSA
 DDIEMARAGALPPPRHASTGAMHPLSGILNSTDYLR
 YRTLSKIPOLITNVDKGPFLASPSDEILIAKIKERT
 HNTKRYTO -> MAIPTEKSRGALQPRAGCRVARAVR
 ISSIVAOE (in isoform 3).
 /ftid=VSP_000970.
 FT VARIANT 186 186 H -> R (IN STRAIN BALB/C).
 FT VARIANT 455 455 A -> T (IN STRAIN BALB/C).
 FT VARIANT 752 752 C -> Y (IN STRAIN BALB/C).
 FT VARIANT 1006 1006 D -> N (IN STRAIN BALB/C).
 FT CONFLICT 111 111 D -> G (IN REF. 3).
 FT CONFLICT 126 126 I -> V (IN REF. 3).
 FT CONFLICT 198 198 V -> A (IN REF. 3).
 FT CONFLICT 214 214 E -> G (IN REF. 3).
 FT CONFLICT 537 537 V -> A (IN REF. 3).
 FT CONFLICT 908 908 S -> P (IN REF. 3).
 FT CONFLICT 929 929 G -> R (IN REF. 3).
 SQ SEQUENCE 1162 AA; 126885 MM; A9455F7F10B61E46 CRC64;
 Query Match 95.4%; Score 5797.5; DB 1; Length 1162;
 Best Local Similarity 96.0%; Pred. No. 1.2e-292;
 Matches 116; Conservative 10; Mismatches 32; Indels 5; Gaps 3;

DB 121 AVIMFILNFEVMEKMGVSPAHDTNHRGPTSMLAGRAKTRFLKIPALLALTARRESSV 180
 181 RSGAGAGAGAPGAVYVDYDLTPAASSESLALDEVYAMDHNHVAAGLPAERATLVPG-- 238
 181 RTGSMHSGAGAPGAVYVDYDLTPAASSESLALDEVYAMDHNHVAAGLPAERATLVPGSA 240
 239 SPSPSARGOLSPRAHSLNDPAGSSCSLARTSRSCASVRASSADDIEMARAGALVP 298
 241 SPVASINGPSPRAQSLNDPAGSSCSLARTSRSCASVRASSADDIEMARAGALVP 300
 299 PPRHASTGAMHPLSGILNSTDYLRRTISKIPQITLNFVDLKDGPLASPTSDREI 358
 301 PPRHASTGAMHPLSGILNSTDYLRRTISKIPQITLNFVDLKDGPLASPTSDREI 360
 359 IAPKIKERTNVTETKVQVLSLGDVLPYKLAQPRIRHWTILHYSPPKAVMDLILLV 418
 361 IAPKIKERTNVTETKVQVLSLGDVLPYKLAQPRIRHWTILHYSPPKAVMDLILLV 420
 419 IYTAVFPTPSAFLKTEEGPPTAECGYACQPLAVVDLYDIMEFTVDILINPTTYVNA 478
 421 IYTAVFPTPSAFLKTEEDGSOAPDCGYACQPLAVVDLYDIMEFTVDILINPTTYVNA 480
 479 NEEVSHRGRIAVHYFKGWEFLDMVAAIPFDLIFGSGSESLGLKTARLLRVYARK 538
 481 NEEVSHRGRIAVHYFKGWEFLDMVAAIPFDLIFGSGSESLGLKTARLLRVYARK 540
 539 LDREYEGAAVFLMCTFALLAHWLACIYVAGNMEDQPHMDSRIGLHNLGDOIGKPVN 598
 541 LDREYEGAAVFLMCTFALLAHWLACIYVAGNMEDQPHMDSRIGLHNLGDOIGKPVN 600
 599 SSGIGGSPSIDKVTALYFTFESSLTSYGFNVSPNTSEKIFSCVWLIGSIMASTFGN 658
 601 SSGIGGSPSIDKVTALYFTFESSLTSYGFNVSPNTSEKIFSCVWLIGSIMASTFGN 660
 659 VSAIIOGLYSGTARVHTOMLRVREIFRHOIPNLRORLEHYOHANSTNGIDMNVLK 718
 661 VSAIIOGLYSGTARVHTOMLRVREIFRHOIPNLRORLEHYOHANSTNGIDMNVLK 720
 719 GFPECLDADICLHNSRLDCHCKPFGATYGCCLALAMKRTTHAPPGDTLVHAGDLLTA 778
 721 GFPECLDADICLHNSRLDCHCKPFGATYGCCLALAMKRTTHAPPGDTLVHAGDLLTA 780
 779 IYFISRSISIELRQDVVVALLGKNDIGEPNLNLYARPKSGNDVRALTYCDLKHITHDDL 838
 781 IYFISRSISIELRQDVVVALLGKNDIGEPNLNLYARPKSGNDVRALTYCDLKHITHDDL 840
 839 LEVLDMPPEFSDHFWSSLEITFNLDTNMIIPGSPGTELEGFESRQRRKLSPFRRTDKD 898
 841 LEVLDMPPEFSDHFWSSLEITFNLDTNMIIPGSPGTELEGFESRQRRKLSPFRRTDKD 900
 899 TEOPGEVSAL--GPRAGAGPSSRGPRGPGWGSPPSSGESSDEDEGGRSSPLRLV 956
 901 TEOPGEVSALGOGGARVPGPSCRGCPGPGWGSPPSSGESSDEDEGGRSSPLRLV 960
 957 PFSSPRPGRPGGEPGLMEDCEKSSDFCNPLSGAFSGVNTFSFGWSRGROYOELRCP 1016
 961 PFSSPRPGRPGGEPGLMEDCEK--SDTCNPLSGAFSGVNTFSFGWSRGROYOELRCP 1019
 1017 APPPSLINPLSPGRPRGDVDSRLDALROLRLRLETRLSADATVYLOLQROMTLVPP 1076
 1020 APPPSLINPLSPGRPRGDVDSRLDALROLRLRLETRLSADATVYLOLQROMTLVPP 1079
 1077 AYSAVTTPGGPTSTSLPLVSPPLTLTDLSSOVQPMACCELPAGAPELPDGPPTRRL 1136
 1080 AYSAVTTPGGPTSTSLPLVSPPLTLTDLTSLSSOVQPMACCELPAGAPELPDGPPTRRL 1139
 1137 SLPGOLGALTSOPLHRHSGDPS 1159
 1140 SLPGOLGALTSOPLHRHSGDPS 1162
 RESULT 5

ID	KCH2_RAT	STANDARD;	PRT;	1163 AA.
AC	008962:	008720:		
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go related gene potassium channel 1) (ERG1) (r-ERG) (RENG) (Ether-a-go-go related protein 1) (Eag related protein 1).			
CN	KCNH2 OR ERG.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
NCBI_TaxID=10116;	[1]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Brain cortex;			
RC	TISSUE=Brain cortex;			
RX	MEDLINE=98329322; PubMed=9664620;			
RA	Bauer C.K., Engeland B., Wulfsen I., Ludwig J., Pongs O.,			
RA	Schwarz J.R.;			
RT	"RENG is a molecular correlate of the inward-rectifying K current in			
RT	clonal net pituitary cells.";			
RL	Recept. Channels 6:19-29(1998).			
RL	[2]			
RP	SEQUENCE OF 409-568 FROM N.A.			
RX	MEDLINE=97164986; PubMed=9012748;			
RA	Wymore R.S., Gintant G.A., Wymore R.T., Dixon J.E., McKinnon D.,			
RA	Cohen I.S.;			
RT	"Tissue and species distribution of mRNA for the IKr-like K+ channel,			
RT	erg."			
RL	Circ. Res. 80:261-268(1997).			
RL	[3]			
RP	TISSUE SPECIFICITY.			
RX	MEDLINE=20183472; PubMed=10718922;			
RA	Wulfsen I., Hauber H.P., Schlemann D., Bauer C.K., Schwarz J.R.;			
RT	"Expression of mRNA for voltage-dependent and inward-rectifying K			
RT	channels in GH3/B6 cells and rat pituitary.";			
RL	J. Neuroendocrinol. 12:263-272(2000).			
RN	[4]			
RP	INTERACTION WITH KCNH6 AND KCNH7, AND MUTAGENESIS OF GLY-630.			
RX	MEDLINE=21079731; PubMed=11212207;			
RA	Wimmers S., Wulfsen I., Bauer C.K., Schwarz J.R.;			
RT	"Erg1, erg2 and erg3 K channel subunits are able to form			
RT	heteromultimers.";			
RL	PLfugers Arch. 441:450-455(2001).			
CC	-1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly			
CC	rectifying potassium channel. Channel properties are modulated by			
CC	cAMP and subunit assembly. Mediates the rapidly activating			
CC	component of the delayed rectifying potassium current in heart			
CC	(IKr) (By similarity).			
CC	-1- SUBUNIT: The potassium channel is probably composed of a homo- or			
CC	heterotetrameric complex of pore-forming alpha subunits that can			
CC	associate with modulating beta subunits. Heteromultimer with			
CC	KCNH6/ERG2 and KCNH7/ERG3. Heteromultimer with			
CC	KCNEL and KCNE2 (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- TISSUE SPECIFICITY: Highly expressed in brain and testis, slightly			
CC	less so in heart, adrenal, retina and thymus. Detected at lower			
CC	levels in lung, soleus, tibialis, and at very low levels in cornea			
CC	and lens. A shorter transcript is detected in skeletal muscle.			
CC	Found in pituitary.			
CC	-1- DOMAIN: The segment S4 is probably the voltage-sensor and is			
CC	characterized by a series of positively charged amino acids at			
CC	every third position.			
CC	-1- PMW: Phosphorylated on serine and threonine residues (By			
CC	similarity).			
CC	-1- SIMILARITY: Belongs to the potassium channel family. H (Eag			
CC	subfamily).			
CC	-1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.			
CC	-1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.			
CC	-1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by aqd for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

[illegible]

Query Match	95.2%; Score 5790; DB 1; Length 1163;
Best Local Similarity	95.6%; Pred. No. 3e-292;
Matches 1112; Conservative 12; Mismatches 35; Indels 4; Gaps 2;	
QY	1 MPVRGHVAPONTFLDTIRKEEGOSKRFTIANARVENCAYICNDGFCCLGYSRAEVM 60
Db	1 MPVRGHVAPONTFLDTIRKEEGOSKRFTIANARVENCAYICNDGFCCLGYSRAEVM 60
QY	61 ORPCTCDFLHGPRFQRRRAAQAIAQALLGAERKEKEIAFYRKDSCFCFLDYVPVKKEDG 120
Db	61 ORPCTCDFLHGPRFQRRRAAQAIAQALLGAERKEKEIAFYRKDSCFCFLDYVPVKKEDG 120
QY	121 AVIMEFLTFEYVMKEKDWGSPAHDTNHRGPTSMLAGRAKTFPLKIPALLALTARSSV 180
Db	121 AVIMEFLTFEYVMKEKDWGSPAHDTNHRGPTSMLAGRAKTFPLKIPALLALTARSSPM 180
QY	181 RSGGAGGAGAGAVVVDVLTLPAPPSSESLALDEVTAMDNDHVAALGPAEERALVGP--G 238
Db	181 RTGTSGSGAGAVVVDVLTLPAPPSSESLALDEVSMNDHNVAGLGAERERRALVGSAA 240
QY	239 SPPRSAPGQLPSPRAHSINLPDASGSSCLATRSRESCASVRRASADDIEMARAGVLP 298
Db	241 SPVASIRPHPSPPAAQSLNPDASGSSCLATRSRESCASVRRASADDIEMARAGALPL 300

QY 299 PPHASTGAMHPLRSGLNSTSDSDLVRRYRTISKIPQITLNFVLDKGPFLASPTSDREI 358
 DB 301 PPHASTGAMHPLRSGLNSTSDSDLVRRYRTISKIPQITLNFVLDKGPFLASPTSDREI 360
 QY 359 IAKIKERTHNVEKTYOVLISGLADVLPEKLOAPRIHRTILHSPFKAWDWMLLLLV 418
 DB 361 IAKIKERTHNVEKTYOVLISGLADVLPEKLOAPRIHRTILHSPFKAWDWMLLLLV 420
 QY 419 IYNAVTPYSAFLKETEKGPAECYACQPLAVVLDVIMFVLDILINFRITYVNA 478
 DB 421 IYNAVTPYSAFLKETEKGPAECYACQPLAVVLDVIMFVLDILINFRITYVNA 480
 QY 479 NEEVSHPGRIAVHYFKGWFLIDMVAIPEDLLIFSGSBEILGLKTARILRYEVAR 538
 DB 481 NEEVSHPGRIAVHYFKGWFLIDMVAIPEDLLIFSGSBEILGLKTARILRYEVAR 540
 QY 539 LDRYSEYGAAYFLMCTPALLIHMVLAICTYAGNNEQPMDSRIGMLHGLQIGKPYN 598
 DB 541 LDRYSEYGAAYFLMCTPALLIHMVLAICTYAGNNEQPMDSRIGMLHGLQIGKPYN 600
 QY 599 SSGIGSPSIKDKVYTAIFYFSSLTIVGFGVNSPNTNSEKIFISCVMLIGSLMYASIFGN 658
 DB 601 SSGIGSPSIKDKVYTAIFYFSSLTIVGFGVNSPNTNSEKIFISCVMLIGSLMYASIFGN 660
 QY 659 VSAIIORLYSGTARYHTQMLRYEFTIRFQIIPNPLRQRLSEYPOHAMSTYNGIDMNAVYK 718
 DB 661 VSAIIORLYSGTARYHTQMLRYEFTIRFQIIPNPLRQRLSEYPOHAMSTYNGIDMNAVYK 720
 QY 719 GPPECQADICHLNLSLLOHCKPFRGATKGCIALAMKFTTHAPPGDTLVHAGDLLTA 778
 DB 721 GPPECQADICHLNLSLLOHCKPFRGATKGCIALAMKFTTHAPPGDTLVHAGDLLTA 780
 QY 779 LVFISGSIIEIRGDDVVAIILGKNDIFGPEPLNYARPKSGNGVRLATYCDLHKIHRLD 838
 DB 781 LVFISGSIIEIRGDDVVAIILGKNDIFGPEPLNYARPKSGNGVRLATYCDLHKIHRLD 840
 QY 839 LEVLDMPYBESDHFMSLITFENLRDTNMIPEGSPGTELEGFSRORRKLSPFRRTDXD 898
 DB 841 LEVLDMPYBESDHFMSLITFENLRDTNMIPEGSPGTELEGFSRORRKLSPFRRTDXD 900
 QY 899 TRQPGVSAAL--GPGNAGAGPSSRGPRGPGWGSPPSSPSSSEDEGGRSSPLRLV 956
 DB 901 TRQPGVSAAL--GPGNAGAGPSSRGPRGPGWGSPPSSPSSSEDEGGRSSPLRLV 960
 QY 957 PFSSPPPGPBGPEGLMEDCEKSDTCNPISGAFSGVNSIFEFWSDSGRQOELPRCP 1016
 DB 961 PFSSPPPGPBGPEGLMEDCEKSDTCNPISGAFSGVNSIFEFWSDSGRQOELPRCP 1020
 QY 1017 APPSLIINTPLSSPGRRPRGVDVSRDLAQRLNRLRLETRLSADMATVQLLOQOMTLVPP 1076
 DB 1021 APPSLIINTPLSSPGRRPRGVDVSRDLAQRLNRLRLETRLSADMATVQLLOQOMTLVPP 1080
 QY 1077 AYSAVTTPRGPRSTSPPLPVSPPLPILTLDSLSQVQFMACEFLPGCAPLPDEGPTTRL 1136
 DB 1081 AYSAVTTPRGPRSTSPPLPVSPPLPILTLDSLSQVQFMACEFLPGCAPLPDEGPTTRL 1140
 QY 1137 SLPGQIGALTSPOLHHRGSDPGS 1159
 DB 1141 SLPGQIGALTSPOLHHRGSDPGS 1163

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98054206; PubMed=9390998;
 RA Shi W., Wymore R.S., Wang H.S., Pan Z., Cohen I.S., McKinnon D.,
 RA Dixon J.E.;
 RT "Identification of two nervous system-specific members of the erg
 RT potassium channel gene family.";
 RL J. Neurosci. 17:9423-9432(1997).
 RN [2]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20183472; PubMed=10718922;
 RA Wolfen I., Hauber H.P., Schlemmer D., Bauer C.K., Schwarz J.R.;
 RT "Expression of mRNA for voltage-dependent and inward-rectifying K
 RT channels in GH3/B6 cells and rat pituitary.";
 RL J. Neuroendocrinol. 12:263-272(2000).
 RN [3]
 RP INTERACTION WITH KCNH2 AND KCNH6, AND MUTAGENESIS OF GLY-480.
 RX MEDLINE=21079731; PubMed=11212207;
 RA Wilmers S., Wulfen I., Bauer C.K., Schwarz J.R.;
 RT "Erg1, erg2 and erg3 K channel subunits are able to form
 RT heteromultimers.";
 RL Pflugers Arch. 441:450-455(2001).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Channel properties may be modulated by
 CC CAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotetrameric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits. Heteromultimer with
 CC KCNH2/ERG1 and KCNH6/ERG2.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Detected in total brain, in superior cervical,
 CC mesenteric and celiac ganglia, and at very low levels in retina.
 CC Found in pituitary.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Erg)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF016191; AAB94741.1; -
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc.-C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF000027; CNMP_binding.1.
 DR Pfam: PF00520; ion_trans.1.
 DR Pfam: PF00785; PAC.1.
 DR SMART: SM00100; CNMP.1.
 DR SMART: SM00086; PAC.1.
 DR SMART: SM00091; PAS.1.
 DR PROSITE: PS50042; CNMP_BINDING.3; 1.
 DR PROSITE: PS50113; PAC_FALSE_NEG.
 DR PROSITE: PS50112; PAS_FALSE_NEG.
 DR Transprot: Ion transport; ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AJ291608; GAG14797.1; -
 DR InterPro: IPR000595; cNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc.C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF000027; cNMP_binding.1.
 DR Pfam: PF00520; Ion_trans.1.
 DR Pfam: PF00785; PAC.1.
 DR SMART; SM00100; cNMP.1.
 DR SMART; SM00086; PAC.1.
 DR PROSITE; PSS0042; cNMP_BINDING_3.1.
 DR PROSITE; PSS0113; PAC; FALSE_NEG.
 DR PROSITE; PSS0112; PAS; FALSE_NEG.
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KM Glycoprotein; Multigene family.

FT DOMAIN 1 412 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 413 433 SEGMENT S1 (POTENTIAL).
 FT TRANSSEM 450 470 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 471 494 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 495 515 SEGMENT S3 (POTENTIAL).
 FT TRANSSEM 522 542 SEGMENT S4 (POTENTIAL).
 FT DOMAIN 543 549 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 550 570 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 615 635 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSSEM 642 662 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 663 1195 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 41 70 PAS.
 FT NP_BIND 92 144 cNMP.
 FT CARBOHYD 745 862 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 1195 AA; 135026 MW; 8DCB9BA0580FFC2 CRC64;

Query Match 54.0%; Score 3283.5; DB 1; Length 1195;
 Best Local Similarity 56.4%; Pred. No. 1.3e-162; Indels 181; Gaps 30;
 Matches 714; Conservative 117; Mismatches 253;

QY 1 MPVARGHVAPONTFLDTIIRKFEQSQRKFTIANARVENCAYICNDGFCELCGYRAEVM 60
 Db 1 MPVARGHVAPONTFLDTIIRKFEQSQRKFTIANARVENCAYICNDGFCELCGYRAEVM 60

QY 61 QRCCTCFELHGRPTQORAAQOIAQALLGAERKVEIAFYRKDSQCFCLVDVVPVKMEDG 120
 Db 61 QRCCTCFELHGRPTQORAAQOIAQALLGAERKVEIAFYRKDSQCFCLVDVVPVKMEDG 120

QY 121 AVIMEFILFEVNEKDWGSPADTNHGRPTSWLARGAKTFELKIPALLATLARSVS 180
 Db 121 AVIMEFILFEVNEKDWGSPADTNHGRPTSWLARGAKTFELKIPALLATLARSVS 180

QY 121 VAMMEFINFEYVDEEKAAP-----ERVNPILPVKIVNKLKGFKPGKGLVLTTRKOSL 175
 Db 121 VAMMEFINFEYVDEEKAAP-----ERVNPILPVKIVNKLKGFKPGKGLVLTTRKOSL 175

QY 181 RSGGAGAGAPGAVVVDVLTLPAPASSESLALDEVMTAMDNHVAAGLPAERRALVGP-- 238
 Db 181 RSGGAGAGAPGAVVVDVLTLPAPASSESLALDEVMTAMDNHVAAGLPAERRALVGP-- 238

QY 176 PQED-----PDVVVID-----SKHSDSYAMKHFKSPTEKESCPSPADDTKALIQSQC 225
 Db 176 PQED-----PDVVVID-----SKHSDSYAMKHFKSPTEKESCPSPADDTKALIQSQC 225

QY 239 SPERSAPGOL-----PSPRANSLNPDASGSSCSIAKTRSRSCASVPRASSADIDEAMRAG 294
 Db 239 SPERSAPGOL-----PSPRANSLNPDASGSSCSIAKTRSRSCASVPRASSADIDEAMRAG 294

QY 226 SPVYNISGPLDHSPPKQMDRLYPDMLQSSSOLTHSRRESLSIRASSVHDIIEGF--S 283
 Db 226 SPVYNISGPLDHSPPKQMDRLYPDMLQSSSOLTHSRRESLSIRASSVHDIIEGF--S 283

QY 295 VLPP-----PPRHAS-----TGAMHPLRSGLSLNTSDSVRYRTTISKIQITLNFVDLK 344
 Db 295 VLPP-----PPRHAS-----TGAMHPLRSGLSLNTSDSVRYRTTISKIQITLNFVDLK 344

QY 284 VHKKNIFRDRHASEDNANRANKGPPNHRKSSLSLGSSTSNLNKSTIKIKIQLTLNPSDK 343
 Db 284 VHKKNIFRDRHASEDNANRANKGPPNHRKSSLSLGSSTSNLNKSTIKIKIQLTLNPSDK 343

QY 345 GDFPLAS-PTSDBEIIAPKIKERTHNVEKVTQVLSGADVPEYKLOAPRIHRWTILHY 403
 Db 345 GDFPLAS-PTSDBEIIAPKIKERTHNVEKVTQVLSGADVPEYKLOAPRIHRWTILHY 403

Db 344 TEKNTSPSSDKTIIPAKYERTHNVEKVTQVLSGADVPEYKLOAPRIHRWTILHY 403
 QY 404 SPFAVMDWILLIVYTAFTPYSAFLEKTEBEGPRAPTECGAACQAPLAVDILVIMF 463
 Db 404 SPFAVMDWILLIVYTAFTPYSAFLEKTEBEGPRAPTECGAACQAPLAVDILVIMF 463

QY 464 IVDLINFRTTYVANEVSHPRIAVHFYFKGFLLDMVAALPFDDLIFGSGSEE--L 520
 Db 464 IVDLINFRTTYVANEVSHPRIAVHFYFKGFLLDMVAALPFDDLIFGSGSEE--L 520

QY 521 IGLLTARLLRLVAVARKLDRESEYGAVALFELMCTPALAHMLACIWAIGNNEOPDM 580
 Db 521 IGLLTARLLRLVAVARKLDRESEYGAVALFELMCTPALAHMLACIWAIGNNEOPDM 580

QY 523 IGLLTARLLRLVAVARKLDRESEYGAVALFELMCTPALAHMLACIWAIGNNEOPDM 582
 Db 523 IGLLTARLLRLVAVARKLDRESEYGAVALFELMCTPALAHMLACIWAIGNNEOPDM 582

QY 581 SRIGMLNLDQIGKPNV-SGGLGPEISIKKQYATLFTSSLSVGFVNSPPTNSEKI 639
 Db 581 SRIGMLNLDQIGKPNV-SGGLGPEISIKKQYATLFTSSLSVGFVNSPPTNSEKI 639

QY 583 DKIGMLDLSLTQIGKRYNDSSSGPIKKQYATLFTSSLSVGFVNSPPTNSEKI 642
 Db 583 DKIGMLDLSLTQIGKRYNDSSSGPIKKQYATLFTSSLSVGFVNSPPTNSEKI 642

QY 640 FSICVMLIGSLMVASIGVNSAIIQRLYSGTARYHTOMLVREFIRHOIINPLRLEE 699
 Db 640 FSICVMLIGSLMVASIGVNSAIIQRLYSGTARYHTOMLVREFIRHOIINPLRLEE 699

QY 643 FSICVMLIGSLMVASIGVNSAIIQRLYSGTARYHTOMLVREFIRHOIINPLRLEE 702
 Db 643 FSICVMLIGSLMVASIGVNSAIIQRLYSGTARYHTOMLVREFIRHOIINPLRLEE 702

QY 700 YFOHAWSTNGIDMNAVALKGPPECLQADICILNRSILOKCPREGATKGLRALAMKFK 759
 Db 700 YFOHAWSTNGIDMNAVALKGPPECLQADICILNRSILOKCPREGATKGLRALAMKFK 759

QY 703 YFOHAWSTNGIDMNAVALKGPPECLQADICILNRSILOKCPREGATKGLRALAMKFK 762
 Db 703 YFOHAWSTNGIDMNAVALKGPPECLQADICILNRSILOKCPREGATKGLRALAMKFK 762

QY 760 TTIAPEDDTLVHAGDILTALYFISRGSEIELRGDVVAALIGKNDIFEPYMLVARPEKSN 819
 Db 760 TTIAPEDDTLVHAGDILTALYFISRGSEIELRGDVVAALIGKNDIFEPYMLVARPEKSN 819

QY 763 TTIAPEDDTLVHAGDILTALYFISRGSEIELRGDVVAALIGKNDIFEPYMLVARPEKSN 822
 Db 763 TTIAPEDDTLVHAGDILTALYFISRGSEIELRGDVVAALIGKNDIFEPYMLVARPEKSN 822

QY 820 GDVRLYCDLHKIHRDLEVLDMYEPEDHFESSLEITPNLRTMIGSGPSTLEG 879
 Db 820 GDVRLYCDLHKIHRDLEVLDMYEPEDHFESSLEITPNLRTMIGSGPSTLEG 879

QY 823 ADVRALTYCDLHKIHRDLEVLDMYEPEDHFESSLEITPNLRTMIGSGPSTLEG 880
 Db 823 ADVRALTYCDLHKIHRDLEVLDMYEPEDHFESSLEITPNLRTMIGSGPSTLEG 880

QY 880 GFGRORRKLFRFRRTDKTEQPEVYALPGRAGAPSS-----RGRPGPWESESPSSGP 935
 Db 880 GFGRORRKLFRFRRTDKTEQPEVYALPGRAGAPSS-----RGRPGPWESESPSSGP 935

QY 881 DTGKLRRLRLSFESEGEKDSKESAN-----DADSDTDTIRRYGSKKHFEKRSRS 933
 Db 881 DTGKLRRLRLSFESEGEKDSKESAN-----DADSDTDTIRRYGSKKHFEKRSRS 933

QY 936 SPSESSDEGPRGSSSLRL-VPPSSPR-----PGEPPGGERLMECKSS 981
 Db 936 SPSESSDEGPRGSSSLRL-VPPSSPR-----PGEPPGGERLMECKSS 981

QY 934 SSFISIDID-----EQKPLFLGTVDSTPRKATRLNGEETMPPSPRIHTKRSHSCHDIT 989
 Db 934 SSFISIDID-----EQKPLFLGTVDSTPRKATRLNGEETMPPSPRIHTKRSHSCHDIT 989

QY 982 DT-----CNPLSGAFSGVSNIFSEFGDSRGROYOLPCRAPTPSLINPL 1027
 Db 982 DT-----CNPLSGAFSGVSNIFSEFGDSRGROYOLPCRAPTPSLINPL 1027

QY 990 DTHSWERPRAPDEECP-----SGLQR--AAMGSE-----TESDLTY-- 1027
 Db 990 DTHSWERPRAPDEECP-----SGLQR--AAMGSE-----TESDLTY-- 1027

QY 1028 \$SPGRPRGDVESRLDALQRLNLETRLSADMATVQLQLOQMTLVPRAYSATTPGPG 1087
 Db 1028 \$SPGRPRGDVESRLDALQRLNLETRLSADMATVQLQLOQMTLVPRAYSATTPGPG 1087

QY 1088 PTSTPPLL-----PVPLP-----TLTL 1105
 Db 1088 PTSTPPLL-----PVPLP-----TLTL 1105

QY 1077 AEYQRPRLRLKTSHPRASIKTRDSFSSQCEFFLDLEKSKLOSKESSLSSGRRLNASE 1136
 Db 1077 AEYQRPRLRLKTSHPRASIKTRDSFSSQCEFFLDLEKSKLOSKESSLSSGRRLNASE 1136

QY 1106 DSLQOV--SOQMACEELPPGA-----PELPQEGPPTRRSLPGLGALTSQPLHRH 1153
 Db 1106 DSLQOV--SOQMACEELPPGA-----PELPQEGPPTRRSLPGLGALTSQPLHRH 1153

QY 1137 DNLFSLKQSDASSELDPQOKTYLHPRIHPSPSS-----LSTVAILG-----LHRH 1186
 Db 1137 DNLFSLKQSDASSELDPQOKTYLHPRIHPSPSS-----LSTVAILG-----LHRH 1186

QY 1154 GSDPG 1158
 Db 1154 GSDPG 1158

QY 1187 VSDPG 1191
 Db 1187 VSDPG 1191

RESULT 8
 KCH7_HUMAN
 ID KCH7_HUMAN STANDARD; PRT; 1196 AA.
 AC ONSA10;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 7 (Ether-a-go-go related gene potassium channel 3) (HERG-3) (Ether-a-go-go related

protein 3) (Eag related protein 3).
 DE KCMH7 OR ERG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue=Brain;
 RA Ganetzy B., Titus S.A.;
 RT "Polynucleotides encoding herg-3 potassium channel."
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1 FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Channel properties may be modulated by cAMP and subunit
 CC assembly.
 CC -1 SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotetrameric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits. Heteromultimer with
 CC KCMH2/ERG1 and KCMH6/ERG2 (By similarity).
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1 DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1 SIMILARITY: Belongs to the potassium channel family. H (Eag)
 CC subfamily.
 CC -1 SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1 SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1 SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF032897; AAD01946.1; -
 DR Genbank: HGMG:18863; KCMH7.
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR00700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR Pfam: PF00785; PAC; 1.
 DR SMART: SM00100; CNMP; 1.
 DR SMART: SM00086; PAC; 1.
 DR SMART: SM00091; PAS; 1.
 DR PROSITE: PS50042; CNMP_BINDING_3; 1.
 DR PROSITE: PS50113; PAC; FALSE_NEG.
 DR PROSITE: PS50112; PAS; FALSE_NEG.
 DR Transport: Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Glycoprotein; Multigene family.
 FT DOMAIN 1 412
 FT TRANSMEM 413 433 SEGMENT S1 (POTENTIAL).
 FT TRANSMEM 450 470 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 471 494 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 495 515 SEGMENT S3 (POTENTIAL).
 FT TRANSMEM 522 542 SEGMENT S4 (POTENTIAL).
 FT DOMAIN 543 549 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 550 570 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 615 635 SEGMENT S6 (POTENTIAL).
 FT TRANSMEM 642 662 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
 FT DOMAIN 663 1196 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 41 70 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 92 144 PAS.
 FT NP_BIND 745 862 CNMP.

FT CAROBYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1196 AA; 135012 MW; 7CE10C31A76DC4FE CRC64;
 Query Match 53.5%; Score 3255; DB 1; Length 1196;
 Best Local Similarity 54.8%; Pred. No. 4e-161;
 Matches 700; Conservative 126; Mismatches 246; Indels 206; Gaps 25;
 QY 1 MPVRGHVAPQNTFLDTIIRKEGOSRKFTIARAVENCAYIYNDGFCGLGYRAYVM 60
 DB 1 MPVRGHVAPQNTFLDTIIRKEGOSRKFTIARAVENCAYIYNDGFCGLGYRAYVM 60
 QY 61 QRPCTCDLHAPRPRORRAAQAQALLGAERKEVIATYRKDGSCFLVDVVPYKNDG 120
 DB 61 QRPCTCDLHAPRPRORRAAQAQALLGAERKEVIATYRKDGSCFLVDVVPYKNDG 120
 QY 121 AVIFILNFEVYMEKDWGSPAHPTNHRGPTSMAPGRATFRKLTPALLALARESSV 180
 DB 121 VAMFPIEFVYTDENAKATP-----EYVNPILPKTYNKKRFGCKFGVLVYRKQSL 175
 QY 181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTAMDNHVAIGLPAERRALVGP-- 238
 DB 176 PQED-----PDVVYID---SKHSDSVAMKHKFSPTKSCSPSEADPTKALIQPKC 225
 QY 239 SPPRSAPQL-----PSPRAHSLNDPASGSSCLARTKRESCASYRRASSADIEAMRG 294
 DB 226 SPLVNIQPLDHSFPRKQMDLYPDMLOSSQSLSHRSRESICSTRASSVHDIEGF--G 283
 QY 295 VLPP-----PPIHRS-----TGAMHPLRSGLNTSDSDLYRYRTISXIPOTILNFVLDK 344
 DB 284 VHPKNIFPRDRASEDNGNVAVGPNNHKKSSLLGSTDNLKYSINKIPDLTLNFSYK 343
 QY 345 GDPFLAS-PTSDRETIAPKIKERTHNTEKYTVYLSLGDVLYPEKKIQAPIHMTILAY 403
 DB 344 TEKKNSPPSSDITIAIPKVKDRHNTEKYTVYLSLGDVLYPEKKIQAPIHMTILAY 403
 QY 404 SPFRVMDMLLLLYIYAVFTPYSAFLKETEBCGACQOLAVVDLYDIME 463
 DB 404 SPFRVMDMLLLLYIYAVFTPYSAFLKETEBCGACQOLAVVDLYDIME 462
 QY 464 IVDILINFRITYVANEEVSHPRIAVHYKGFLLDMVAIFDILLIFSGSEE---L 520
 DB 463 IVDILINFRITYVANEEVSHPRIAVHYKGFLLDMVAIFDILLIFSGSEE---L 522
 QY 521 IGLKTAFLRLVAVARKIDRYSEYGAAYFLMCTALIAHMLACIYAIGNEDQPHMD 580
 DB 523 IGLKTAFLRLVAVARKIDRYSEYGAAYFLMCTALIAHMLACIYAIGNEDQPHMD 582
 QY 581 SRIGMHNHLDGQIGKRPYV-SSGLGPRSTKDKYVVALYFTFSSLSVGVGNVSPNTNISKI 639
 DB 583 DKIGMLDLSLGGQIGKRIYNDSSSSGSPSTKDKYVVALYFTFSSLSVGVGNVSPNTNISKI 642
 QY 640 FSIQVMLIGSLMAYASIFGNVSAITQRLYSGTARHTQQLRVREFIRFQIPIPRORLEE 699
 DB 643 FSIQVMLIGSLMAYASIFGNVSAITQRLYSGTARHTQQLRVREFIRFQIPIPRORLEE 702
 QY 700 YFOHAWSYTYNGIDMNAVLYKGFPECLQADICHLNLSLQNHCKPRPGATKGCRLAMKFK 759
 DB 703 YFOHAWSYTYNGIDMNAVLYKGFPECLQADICHLNLSLQNHCKPRPGATKGCRLAMKFK 762
 QY 760 TTHAPRPGDTLVHADDLTALYFIRGSEILRGVYVAIIGKNIDFEPPLVYARPKSN 819
 DB 763 TTHAPRPGDTLVHADDLTALYFIRGSEILRGVYVAIIGKNIDFEPPLVYARPKSN 822
 QY 820 GDVALYLYCDLHKIHRDILLVLMYEPESHPFSSLEITPNLLDTJNM---IPSPSTE 876
 DB 823 ADVALLYLYCDLHKIHRDILLVLMYEPESHPFSSLEITPNLLDTJNM---IPSPSTE 882
 QY 877 LEGGFSRQRRKRLSFRRTDKDTPQPEVSAIGRAGAGPSSNGRPGGPGESPPSSGPPS 936
 DB 883 SEGQNCRLRRKRLSFESEGEKE-----EGPGRSSPLRLVFPSSPPGPGGPIMEDCEKSSDT 983
 QY 937 SPESSED-----EGPGRSSPLRLVFPSSPPGPGGPIMEDCEKSSDT 983

Db 909 DPEDSADITRNHYOSKRHEEKKSRSSSTISSI-----DDE 944
 Qy .984 CNLGSAGFSGVSNIFSPWGDSDRGROYELRCPAPPSLNLIPL----- 1027
 Db 945 QKPL---FSGIYVSSPEIGKASGLDFEE---TVPISGRNHDIKRSHSCDITDMKSWER 997
 Qy 1028 -----SSPGRRRP-----GDVESRDLAQOLNLETRLASADMATVL 1064
 Db 998 ENAHPOEDSSPALGRAAMGISETESDLYGEVEQLDLDLQOLNLESGQMTDITL 1057
 Qy 1065 QLLQKQKTLVPPAYSAVTTPGPPPTSPPL-----PVSPLEPT-LTLDLSQVQFMA 1116
 Db 1058 QLLQKQKTLVPPAYSAVTTPGPPPTSPPL-----PVSPLEPT-LTLDLSQVQFMA 1114
 Qy 1117 CEELPPCAPLPQEG-----PTBRLSLP-G 1140
 Db 1115 LEKSKLKSRESLSSGVLNLTASDNLTLKQSDLSLELHAKRKTYVPIRHPSLPDS 1174
 Qy 1141 QILGALTSPQLRHGSDPG 1158
 Db 1175 SLSTVGIVGLRHRYSDPG 1192
 RESULT 9
 KCH6_HUMAN STANDARD: PRT: 994 AA.
 ID KCH6_HUMAN 09H252; 09BRD7;
 AC 09H252; 09BRD7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 6 (Ether-a-go-go related gene potassium channel 2) (Ether-a-go-go related protein 2)
 DE (Eag related protein 2).
 GN KCHN6 OR ERG2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Titus S.A., Ganetzky B.S.,
 RT "Human Eag-related gene member 2 (Herg2) potassium channel,"
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC Tissue-Amygdala, and Kidney;
 RA Minomiyu K., Magatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H., Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hiraio M., Shimizu F., Wakebe H., Oho T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RT "HMO human cDNA sequencing project,"
 RL Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC Tissue-Oterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Casanovi P., Prange C., Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarsson P.H., Richards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., RA Roderfeld Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E., RA Schmeckel A., Schein J.E., Jones S.J.M., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium channel. Elicits a slowly activating, rectifying current (by similarity). Channel properties may be modulated by cAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or heterotetrameric complex of pore-forming alpha subunits that can associate with modulating beta subunits. Heteromultimers with KCNH2/ERG1 and KCNH7/ERG3 (by similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Experimental confirmation may be lacking for some isoforms;
 CC Name=1;
 CC IsoId=Q9H252-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9H252-2; Sequence=VSP_000977, VSP_000978;
 CC Name=3;
 CC IsoId=Q9H252-3; Sequence=VSP_000979, VSP_000980;
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is characterized by a series of positively charged amino acids at every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag) subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF311913; AAC40871.1; -;
 CC EMBL: AK090969; BAC03559.1; -;
 CC EMBL: AK091877; BAC03764.1; -;
 CC EMBL: BC006334; AA06334.1; -;
 CC Genew: HCN6; KCHN6.
 CC InterPro: IPR000595; CNMP_binding.
 CC InterPro: IPR003967; Erg_channel.
 CC InterPro: IPR005821; Ion_trans.
 CC InterPro: IPR001622; K+channel_pore.
 CC InterPro: IPR005820; M+channel_nlg.
 CC InterPro: IPR001610; PAC.
 CC InterPro: IPR000700; PAS-assoc.C.
 CC InterPro: IPR000014; PAS_domain.
 CC Pfam: PF00027; CNMP_binding; 1.
 CC Pfam: PF00520; Ion_trans; 1.
 CC Pfam: PF00785; PAC; 1.
 CC SMART: SM00100; CNMP; 1.
 CC SMART: SM00086; PAC; 1.
 CC PROSITE: PSS0042; CNMP_BINDING_3; 1.
 CC PROSITE: PSS0115; PAC; FALSE_NEG.
 CC PROSITE: PSS0112; PAS; FALSE_NEG.
 CC Transprot: Ion transport; Ionic channel; Voltage-gated channel;
 CC Potassium channel; Potassium; Potassium transport; Transmembrane;
 CC Multigene family; Alternative splicing.
 CC DOMAIN 1 261 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 262 282 SEGMENT S1 (POTENTIAL).
 CC TRANSMEM 299 319 SEGMENT S2 (POTENTIAL).
 CC DOMAIN 320 340 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 341 361 SEGMENT S3 (POTENTIAL).
 FT TRANSMEM 371 391 SEGMENT S4 (POTENTIAL).
 FT DOMAIN 392 398 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 399 419 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 464 484 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSMEM 491 511 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 512 994 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 41 70 PAS.
 FT DOMAIN 92 144 PAC.
 FT NP_BIND 594 711 CNMP.
 FT VASPLIC 419 472 MYATGNVERPLYLEHKIGWLDISGVOLGKRYNGSDPASPSV
 ODKYVTALETFPS -> C (in isoform 2).
 FT VASPLIC 745 780 /FtId=VSP_000917.
 FT VASPLIC 501 502 Missing (in isoform 2).
 FT VASPLIC 503 994 SL -> CE (in isoform 3).
 FT VASPLIC 503 994 /FtId=VSP_000979.
 FT VASPLIC 503 994 Missing (in isoform 3).
 FT CONFLICT 925 925 T -> M (in REF. 2; BAC03764).
 FT CONFLICT 963 963 F -> L (in REF. 2; BAC03764).
 FT SEQUENCE 994 AA; 109924 MM; BE9ECB349A798576 CRC64;

Query Match 50.1%; Score 3043.5; DB: 1; Length 994;
 Best Local Similarity 54.2%; Pred. No. 2.8e-150;
 Matches 670; Conservative 85; Mismatches 153; Indels 329; Gaps 32;

OY 1 MPVRGHAAPONTFLDTIRKFEQSRKFIITANARVENCAYICNDGFCGLGYSRALVM 60
 DB 1 MPVRGHAAPONTYIDTIRKFEQSRKFIITANAMEMCAIITCNDGFCGLGYSRALVM 60
 OY 61 ORPCTCDFLGRPRQRRRAAOIQAALLGAERKVEIAFYRKDGSCFCLVDVVPYKNDG 120
 DB 61 OORPCDFLGRPRPSSAVSRALQALLGAEECKVDILEYRKDASSFRCGLVDVVPYKNDG 120
 OY 121 AVIMFLNEFYVMKMDVGSRAHDTNNHGRPTSMAPRATKFRKLTPALLA-LTARSS 179
 DB 121 AVIMFLNEFYVMKMDVGSRAHDTNNHGRPTSMAPRATKFRKLTPALLA-LTARSS 179
 OY 121 AVIMFLNEFYVMKMDVGSRAHDTNNHGRPTSMAPRATKFRKLTPALLA-LTARSS 179
 DB 121 AVIMFLNEFYVMKMDVGSRAHDTNNHGRPTSMAPRATKFRKLTPALLA-LTARSS 179
 OY 180 VR-----SGGAGAGAPAVVVDVLTTPAAPSSESLALDEVTAMDNHVGAGLPAEERRA 233
 DB 146 QRLSQSFLGSESGHGRG----- 164
 OY 234 LVPGSPRSPAPGOLPSPRAHSLNDASGSSCSLTARTSRBSCASVRRASSADIEAMRA 293
 DB 165 --GPG-----PG----- 169
 OY 294 GVLPRPRHASTGAMHPLRSGLNSTSDSLVRYRTKIQITLNFYDLKGDPRFLASPT 353
 DB 170 -----TGR-----GKYRTISQIDPQPTLNFVEFNLKHRSSST 201
 OY 354 SDRELIAP-KIKERTHNTKVTQVLSGADVLPREKQLQAPRIHRTILHYSPEKAVMD 412
 DB 202 TEIETIAHKYVERQNTKVTQVLSGADVLPREKQLQAPRIHRTILHYSPEKAVMD 261
 OY 413 LILLVITYAVTPYSAFLKETEGRPATECGACOPPLAVVDLIVIMFVILINFR 472
 DB 262 LILLVITYAVTPYSAFLKETEGRPATECGACOPPLAVVDLIVIMFVILINFR 320
 OY 473 TTYVANEEVWSHPRIAVHYFKGFLDMVAALPFDDLIRRTGSDETTTILGLKTKARL 529
 DB 321 TTYVANEEVWSHPRIAVHYFKGFLDMVAALPFDDLIRRTGSDETTTILGLKTKARL 380
 OY 530 LRLVAVARKLDYRSEYGAVALFLMCTPALLIAHMLACIWIYAGNNEOPHMDSRIGMLNL 589
 DB 381 LRLVAVARKLDYRSEYGAVALFLMCTPALLIAHMLACIWIYAGNNEOPHMDSRIGMLNL 440
 OY 590 GDQIGKPYNSG-LGSPKIDKYYTALYTFSSLTSGFGVNSPNTSEKIFPSICVMLIG 648
 DB 441 GVOJGKRYNGSDPASPSVODKYYTALYTFSSLTSGFGVNSPNTSEKIFPSICVMLIG 500
 OY 649 SLMASTIGNSATIORLYSGTARYHQMRLVREIRRHQIPNPKOLEEYFOHANSYT 708
 DB 649 SLMASTIGNSATIORLYSGTARYHQMRLVREIRRHQIPNPKOLEEYFOHANSYT 708

DB 501 SLMASTIGNSATIORLYSGTARYHQMRLVREIRRHQIPNPKOLEEYFOHANSYT 560
 OY 709 NGIDMNAVLKGFPECCADICLHNRSLDCKPFGATKGCRLARANKFTTAPPGDT 768
 DB 561 NGIDMNAVLKGFPECCADICLHNRSLDCKPFGATKGCRLARANKFTTAPPGDT 620
 OY 769 LVHAGDLTALYFISKGIETLRDGVYVAIIGKNDIFEEPILNLYARPKSGNDVALTYC 828
 DB 621 LVHAGDLTALYFISKGIETLRDGVYVAIIGKNDIFEEPILNLYARPKSGNDVALTYC 680
 OY 829 DLHKIHRDLEVLDMYEPFSDHFMSSLEITFENRDT-----NMIPGSGSTELEGFSR 883
 DB 681 DLHKIHRDLEVLDMYEPFSDHFMSSLEITFENRDT-----NMIPGSGSTELEGFSR 738
 OY 884 ORKRLSPRRRTDKDTEOPGEVSALGPGRAGAPSSRPGCPGMESSPSGPS--SPSS 941
 DB 739 -----LSDNGSGSPHE---LGP-----QFPSKGYSLILGPGSQ 767
 OY 942 EDEGPRSSPLRLVFPSSPPRPGEPGCEFLMEDCKSSDTCNPLSGAFSGVNSIFSFW 1001
 DB 768 NSMGAG-----PCAPGHDAAPPL-----SISDASGLW 795
 OY 1002 GDSRGROYQELPCPAPPTSLNIPLSSPGRPRGDE-----SRDLALOROLNRET 1054
 DB 796 PE-----LLQEMF-----PRHSP-OSPOEDPCWPLKLSRLEQLOAQMNRLS 838
 OY 1055 RLSDMATVQLQLORQ-----TLVPAYSAVTPGP-----GPTS 1090
 DB 839 RVSSDLSTILQLQKPMQGHASYILFAPASNDLALVPIA-SETTSPGRLPGFLPPAQ 897
 OY 1091 T-----SPLPVPSPPLTLTDSLSQVQFMACEFLPG-----APLPL 1128
 DB 898 TPYSQDLDCCSPKHNSSPRMP--HLAVATKTLAPSS--OEOPGLMPLASPLRP 951
 OY 1129 OE-----GPTRRLSLPQGLALTSQ-PLHRRSGDPG 1158
 DB 952 LEVQGLICPGCFPS-SLPEHLGSVPRQDLDFQRHSGDPG 987

RESULT 10
 KCH6_RAT
 ID KCH6_RAT STANDARD: PRT: 950 AA.
 AC 054853;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 6 (Ether-a-go-go related gene potassium channel 2) (Ether-a-go-go related protein 2)
 DE (Eeg related protein 2).
 DE KCNH6 OR ERG2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervical ganglion;
 RX MEDLINE=98054206; PubMed=9390998;
 RA Shi W., Wyomere R.S., Wang H.-S., Pan Z., Cohen I.S., McKinnon D.,
 RA Dixon J.E.;
 RT *Identification of two nervous system-specific members of the erg
 RT potassium channel gene family.*;
 RL J. Neurosci. 17:9423-9432(1997).
 RN [2]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20183472; PubMed=10718922;
 RA Wulfsen I., Hauber H.P., Schiemann D., Bauer C.K., Schwarz J.R.;
 RT "Expression of mRNA for voltage-dependent and inward-rectifying K
 RT channels in GH3/B6 cells and rat pituitary.";
 RL J. Neuroendocrinol. 12:263-272(2000).
 RN [3]
 RP INTERACTION WITH KCNH2 AND KCNH7, AND MUTAGENESIS OF GLY-480.
 RX MEDLINE=21079731; PubMed=11212207;

RA Wimmers S., Wulfsen I., Bauer C.K., Schwarz J.R.;
 RT "Erg1, erg2 and erg3 K channel subunits are able to form
 RT heteromultimers.";
 RL Pfugers Arch. 441:450-455(2001).
 RN [4]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=21319165; PubMed=11425889;
 RA Sagalich M.J., Machado E., Rudy B.;
 RT "Differential expression of genes encoding subthreshold-operating
 RT voltage-gated K⁺ channels in brain.";
 RL J. Neurosci. 21:4609-4624(2001).
 CC
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Elicits a slowly activating, rectifying current. Channel
 CC properties may be modulated by cAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotrimeric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits. Heteromultimer with
 CC KCHN2/ERG1 and KCNH7/ERG3.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Highly expressed in celiac and superior
 CC mesenteric ganglia, but not detected in brain or in heart.
 CC detected at low levels in retina. Also found in pituitary.
 CC According Ref.4 found in olfactory bulb (granular and mitral cell
 CC layers).
 CC -1- DOMAIN: The segment s4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL, AF016192; AB94742.1; -
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; ERG_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_intg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR Pfam: PF00785; PAC; 1.
 DR SMART: SM00100; CNMP; 1.
 DR SMART: SM00086; PAC; 1.
 DR PROSITE: PS00042; CNMP_BINDING_3; 1.
 DR PROSITE: PS00113; PAC; FALSE_NEG.
 DR PROSITE: PS00112; PAS; FALSE_NEG.
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Multigene family.
 FT DOMAIN 1 261 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 262 282 SEGMENT S1 (POTENTIAL).
 FT TRANSSEM 299 319 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 320 340 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 341 361 SEGMENT S3 (POTENTIAL).
 FT TRANSSEM 371 391 SEGMENT S4 (POTENTIAL).
 FT DOMAIN 392 398 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 399 419 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 464 484 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSSEM 491 511 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 512 950 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 41 70 PAS.

FT DOMAIN 92 144 PAC.
 FT NP_BIND 594 711 CNMP.
 FT MUTAGEN 480 480 G->S: DOMINANT NEGATIVE MUTANT; ABOLISHES
 FT SEQUENCE 950 AA: 105705 MW: 455A674B08776F0 CRC64;
 SO
 Query Match 49.1%; Score 2982; DB 1; Length 950;
 Best Local Similarity 52.9%; Pred. No. 4e-147;
 Matches 647; Conservative 87; Mismatches 146; Indels 344; Gaps 25;
 QY 1 MPVRGHAAPONTFLDTITIKFEQSKFIITANAVNCAYITNDGFCICGSRAEVM 60
 DB 1 MPVRGHVAPONTFLDTITIKFEQSKFIITANAVNCAYITNDGFCICGSRAEVM 60
 QY 61 ORPCTGFLGPPRTGRRAAAOIAQALGAERKVEIFVRRDSCFCLVDVVPVKEDM 120
 DB 61 ORPCTGFLGPPRTGRRAAAOIAQALGAERKVEIFVRRDSCFCLVDVVPVKEDM 120
 QY 61 ORPCTGFLGPPRTGRRAAAOIAQALGAERKVEIFVRRDSCFCLVDVVPVKEDM 120
 DB 61 ORPCTGFLGPPRTGRRAAAOIAQALGAERKVEIFVRRDSCFCLVDVVPVKEDM 120
 QY 121 AVIMEILNFVMEKMDVGPADHNRGPTSMABGRKTERLKLPAALLATARRSSV 180
 DB 121 AVIMEILNFVMEKMDVGPADHNRGPTSMABGRKTERLKLPAALLATARRSSV 180
 QY 121 AVIMEILNFVMEKMDVGPADHNRGPTSMABGRKTERLKLPAALLATARRSSV 180
 DB 121 AVIMEILNFVMEKMDVGPADHNRGPTSMABGRKTERLKLPAALLATARRSSV 180
 QY 155 -GEGGSHSRP-----DLAQLAKSS-----RSLLORLSHSFL----- 154
 DB 155 -GEGGSHSRP-----DLAQLAKSS-----RSLLORLSHSFL----- 154
 QY 241 PRSAPGQLPSRAHSLNPASGSSCLARTSRSCASVRRASSADDIEAMRAGVLP 300
 DB 241 PRSAPGQLPSRAHSLNPASGSSCLARTSRSCASVRRASSADDIEAMRAGVLP 300
 QY 174 ----- 173
 DB 174 ----- 173
 QY 301 RHAATGAMHPLRSLGSLNSTSDSLVRYRTISKIPQTLNFVDLKGDFLSPDSREIT 360
 DB 301 RHAATGAMHPLRSLGSLNSTSDSLVRYRTISKIPQTLNFVDLKGDFLSPDSREIT 360
 QY 174 ----- 208
 DB 174 ----- 208
 QY 361 P-KIKERTHWTEKVTQVLSGADVLEPKYKQARIRHRTILHSPKAWMDLILLY 419
 DB 361 P-KIKERTHWTEKVTQVLSGADVLEPKYKQARIRHRTILHSPKAWMDLILLY 419
 QY 209 PHKVEYRQNTKVTQVLSGADVLEPKYKQARIRHRTILHSPKAWMDLILLY 268
 DB 209 PHKVEYRQNTKVTQVLSGADVLEPKYKQARIRHRTILHSPKAWMDLILLY 268
 QY 420 YTAVFRTYSAFLKTEEGPATECGYACOPLVNVDIYDIMPVILIRFTTYNAN 479
 DB 420 YTAVFRTYSAFLKTEEGPATECGYACOPLVNVDIYDIMPVILIRFTTYNAN 479
 QY 269 YTAVFRTYSAFLKTEEGPATECGYACOPLVNVDIYDIMPVILIRFTTYNAN 327
 DB 269 YTAVFRTYSAFLKTEEGPATECGYACOPLVNVDIYDIMPVILIRFTTYNAN 327
 QY 480 EEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGSE--DIGELKTARLLRYVA 536
 DB 480 EEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGSE--DIGELKTARLLRYVA 536
 QY 328 DEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGSE--DIGELKTARLLRYVA 387
 DB 328 DEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGSE--DIGELKTARLLRYVA 387
 QY 537 RKLDRYSEGAAYFLMCTFPALIAHMLACIWAIGMEOPHMDSRIGMLHNLDOQK 596
 DB 537 RKLDRYSEGAAYFLMCTFPALIAHMLACIWAIGMEOPHMDSRIGMLHNLDOQK 596
 QY 388 RKLDRYSEGAAYFLMCTFPALIAHMLACIWAIGMEOPHMDSRIGMLHNLDOQK 447
 DB 388 RKLDRYSEGAAYFLMCTFPALIAHMLACIWAIGMEOPHMDSRIGMLHNLDOQK 447
 QY 597 YNSGSG-LGSPRIKRYTALYFPFSSLSVSGFVNSPNTSEKIFSCVMLIGSLMYASI 655
 DB 597 YNSGSG-LGSPRIKRYTALYFPFSSLSVSGFVNSPNTSEKIFSCVMLIGSLMYASI 655
 QY 448 YNSGSDPASGSPVQKRYTALYFPFSSLSVSGFVNSPNTSEKIFSCVMLIGSLMYASI 507
 DB 448 YNSGSDPASGSPVQKRYTALYFPFSSLSVSGFVNSPNTSEKIFSCVMLIGSLMYASI 507
 QY 656 FGNVSATIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEEFQHAMSYTNGIDMNA 715
 DB 656 FGNVSATIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEEFQHAMSYTNGIDMNA 715
 QY 508 FGNVSATIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEEFQHAMSYTNGIDMNA 567
 DB 508 FGNVSATIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEEFQHAMSYTNGIDMNA 567
 QY 716 VLKGFPECLQADICLHLNRSLLOHCKPRGATKCLRLAKMKFTTHAPRGDTLVHAGDL 775
 DB 716 VLKGFPECLQADICLHLNRSLLOHCKPRGATKCLRLAKMKFTTHAPRGDTLVHAGDL 775
 QY 568 VLKGFPECLQADICLHLNRSLLOHCKPRGATKCLRLAKMKFTTHAPRGDTLVHAGDL 627
 DB 568 VLKGFPECLQADICLHLNRSLLOHCKPRGATKCLRLAKMKFTTHAPRGDTLVHAGDL 627
 QY 776 LTAIFYTSRGSIELLRDGVVVAIIIGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 835
 DB 776 LTAIFYTSRGSIELLRDGVVVAIIIGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 835
 QY 628 LSTLYFTSRGSIELLRDGVVVAIIIGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 687
 DB 628 LSTLYFTSRGSIELLRDGVVVAIIIGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 687
 QY 836 DLLLEVLDVMEYFSDHFWSSLEITFNLRDTN---MTP-GSPGTELEGFSRQRRKLS 890
 DB 836 DLLLEVLDVMEYFSDHFWSSLEITFNLRDTN---MTP-GSPGTELEGFSRQRRKLS 890
 QY 688 ADLLEVDVDMYAFADFTMNKLEVFNFNLADNGGLQSPRQAGHQDPQGEF----- 738
 DB 688 ADLLEVDVDMYAFADFTMNKLEVFNFNLADNGGLQSPRQAGHQDPQGEF----- 738
 QY 891 FRRRTDKDTEQPGVYSALGPGACAGPSSRRCGPGMGESPSSGPPSPESSEDEGPGRSS 950
 DB 891 FRRRTDKDTEQPGVYSALGPGACAGPSSRRCGPGMGESPSSGPPSPESSEDEGPGRSS 950
 QY 739 -----LNDGSGAAPS----- 749
 DB 739 -----LNDGSGAAPS----- 749


```

QY 991 SPRLVPPSSPPPEPGEPLMEDCEKSDTCNPLSGAFSGVNIISFWDGSRGXQ 1010
Db 750 -----TISDTSLAM----- 759
QY 1011 ELPRCAPPSLL-NIPSSPGRPRGDVES-----RLDLOLOLNLEFRLSADMT 1062
Db 760 -----PELLQMPSPSP--NPRQDDCWRELGRLEQLOLOQOMNLESRSVSSLSR 808
QY 1063 VLQQLRQMTLVPAY-----SAVTTTPGPGP-----TSTSPLL----- 1095
Db 809 ILQLLQHPQG--RPSYITIGASASDLASPEPESVTRSSSTLLVGHVPSAQTLSYDLD 866
QY 1096 -----PVSPLPTLTLDLSQVQFMACEELPPAPAEPLQEGPPTRLSLPGQLA-- 1144
Db 867 HIGPRNFSPPTPHVAMADKTLVPS-----EQEQPGGLSPLASPLRLPLEVPGIGSGSRF 922
QY 1145 -----LTSOP-----LHRRGSDPG 1158
Db 923 PSLPEHLSSVPKQLEFQRHSDPG 946

RESULT 11
KCH2_CHICK
ID KCH2_CHICK STANDARD: PRT: 526 AA.
AC Q9P784:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go related protein 1) (Bag related protein 1) (Fragment).
GN KCNH2 OR ERG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=20368176; PubMed=10906470;
RA Crociata O., Cherubini A., Piccini E., Polvani S., Costa L.,
RA Fontana L., Hofmann G., Rosati B., Wanke E., Olivetto M.,
RA Arcangeli A.;
RA "ery gene(s)" expression during development of the nervous and muscular
RT system of quail embryos."
RL Mech. Dev. 95:239-243(2000).
CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly
CC rectifying potassium channel (By similarity).
CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
CC heterotetrameric complex of pore-forming alpha subunits that can
CC associate with modulating beta subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -1- SIMILARITY: Belongs to the potassium channel family. H (Bag)
CC subfamily.
CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: AJ271210; CAB6135.1;
CC InterPro: IPR000595; CNMP_binding.
CC InterPro: IPR003967; Erg_channel.
CC InterPro: IPR005821; Ion_trans.

```

```

DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR Pfam: PF00027; CNMP_binding.1.
DR Pfam: PF00520; Ion_trans.1.
DR SMART: SM00100; CNMP.1.
DR PROSITE: PS00042; CNMP_BINDING_3; 1.
KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW Potassium channel; Potassium; Potassium transport; Transmembrane;
KW Glycoprotein; Multigene family.
FT NON_TER 1
FT DOMAIN 1 103 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 104 124 SEGMENT S1 (POTENTIAL).
FT TRANSMEM 151 171 SEGMENT S2 (POTENTIAL).
FT DOMAIN 172 192 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 193 213 SEGMENT S3 (POTENTIAL).
FT TRANSMEM 224 244 SEGMENT S4 (POTENTIAL).
FT DOMAIN 245 250 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 251 271 SEGMENT S5 (POTENTIAL).
FT TRANSMEM 316 336 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT TRANSMEM 343 363 SEGMENT S6 (POTENTIAL).
FT DOMAIN 364 526 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 446 >526
FT CARBOHYD 144 144
FT CARBOHYD 301 301
FT NON_TER 526
SQ SEQUENCE 526 AA; 59793 MW; 2C4D6DBDA9E015C CRC64;

Query Match 37.4%; Score 2275.5; DB 1; Length 526;
Best Local Similarity 82.6%; Pred No. 7.1e-111;
Matches 440; Conservative 33; Mismatches 39; Indels 21; Gaps 4;

QY 295 VLPPPRHASTGAMHPLRSLGSLNSTSDLVYRTISKIPQITLNFVLDKGPFLASPTS 354
Db 10 VKPNP-----NSTSDSLMKYRTISOIPQTLNFEVFNLEKRRSGSTT 53
QY 355 DREITAP-KIKERTHNTEKYQVLSGADVLEPKLQAPRHRHTLIHSPFKVWML 413
Db 54 EELIAPKHYERTONTEKYQVLSGADVLEPKLQAPRHRHTLIHSPFKVWML 113
QY 414 ILLVITYAVFTPYSAALFLKTEEGPPATEGCVACOPLAUVLDLVDIMFVILINPT 473
Db 114 ILLVITYAVFTPYSAALFLKTEEGPPATEGCVACOPLAUVLDLVDIMFVILINPT 173
QY 474 TYVNAEEVSHPGRIAVHYFKGFWLDMVAALIPDLIFSGSGEE--LIGLLTARLL 530
Db 174 TYVNAEEVSHPGRIAVHYFKGFWLDMVAALIPDLIFSGSGEE--LIGLLTARLL 233
QY 531 RLVRARRLDRSEYGAIVFLMCTFALIAHWLACIYVAGNMOPHDSRIGLHWNLG 590
Db 234 RLVRARRLDRSEYGAIVFLMCTFALIAHWLACIYVAGNMOPHDSRIGLHWNLG 293
QY 591 DOIGKPNSSGL-GGPSIKDKYVVALYFTFSSITSVGCVNPSPTNSKIFISICVMLIGS 649
Db 294 DOIGKPNSSGL-GGPSIKDKYVVALYFTFSSITSVGCVNPSPTNSKIFISICVMLIGS 353
QY 650 LMTASIFGVNSAIIORLYSGARHYTOMLRVREIRHQIPPLRQRLLEYFQHAWSYN 709
Db 354 LMTASIFGVNSAIIORLYSGARHYTOMLRVREIRHQIPPLRQRLLEYFQHAWSYN 413
QY 710 GIDMNAVLKGPEDICADICHLNRLSLLOHCKPFGAGKGCRLAMKFKTHAPPGDTL 769
Db 414 GIDMNAVLKGPEDICADICHLNRLSLLOHCKPFGAGKGCRLAMKFKTHAPPGDTL 473
QY 770 VHAGDLTALYFISGSIIEILRGDVVAAILKNDNIFGEPMLLYAPRGSNGDV 822
Db 474 VHAGDLTALYFISGSIIEILRGDVVAAILKNDNIFGEPMLLYAPRGSNGDV 526

RESULT 12
KCH4_HUMAN
ID KCH4_HUMAN STANDARD: PRT: 1017 AA.
AC Q9U005:
DT 28-FEB-2003 (Rel. 41, Created)

```


DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Potassium voltage-gated channel subfamily H member 4 (Ether-a-go-go-
 DE like potassium channel 1) (Elk channel 1) (ELK1) (Brain-specific eag-
 DE like channel 2) (BRC2).
 GN KCNH4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=9386988; PubMed=10455180;
 RA Miyake A., Mochizuki S., Yokoi H., Kohda M., Furuchi K.;
 RT "New ether-a-go-go K⁺ channel family members localized in human
 RT telencephalon.";
 RL J. Biol. Chem. 274:25018-25025(1999).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Elicits an outward current, but shows no inactivation.
 CC Channel properties may be modulated by cAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotetrameric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Detected only in brain, in particular in the
 CC telencephalon. Detected in putamen and caudate nucleus, and at
 CC lower levels in cerebral cortex, occipital and hippocampus.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL, AB022698; BA83592.1; -
 DR Genew; HGNC:6253; KCNH4.
 DR MM; 604528; -
 DR GO: GO:0008076; C:voltage-gated potassium channel complex; TAS.
 DR GO: GO:0005249; F:voltage-gated potassium channel activity; TAS.
 DR GO: GO:0006813; P:potassium ion transport; TAS.
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K_channel_pore.
 DR InterPro: IPR005820; M_channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR Pfam: PF00785; PAC; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00100; CNMP; 1.
 DR SMART; SM00091; PAS; 1.
 DR TIGRFAMs; TIGR00229; sensory_box; 1.
 DR TIGRFAMs; TIGR0042; CNMP_BINDING_3; 1.
 DR PROSITE; PS00112; PAC; 1.
 DR PROSITE; PS00113; PAC; 1.
 KM Transport: Ion transport; Ionic channel; Voltage-gated channel;
 KM Potassium channel; Potassium; Potassium transport; Transmembrane;
 KM Glycoprotein; Multigene family.
 FT DOMAIN 1 228 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 229 249 SEGMENT S1 (POTENTIAL).

FT TRANSMEM 260 280 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 281 302 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 303 323 SEGMENT S3 (POTENTIAL).
 FT TRANSMEM 333 353 SEGMENT S4 (POTENTIAL).
 FT DOMAIN 354 361 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 362 382 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 428 448 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSMEM 483 503 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 504 1017 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 14 90 PAS.
 FT DOMAIN 93 145 CNMP.
 FT NP_BIND 556 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 326 326 CARBOHYD (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1017 AA; 111692 MW; 48480DFC75810A38 CRC64;
 Query Match 26.9%; Score 1636; DB 1; Length 1017;
 Best Local Similarity 34.9%; Pred. No. 1.8e-77;
 Matches 424; Conservative 153; Mismatches 319; Indels 318; Gaps 41;
 QY 1 MPVRGVAQONTFTDITIRKFEQSRRFIITANAR-VENCNAVITCNDGFCGLCYSAEV 59
 D 1 MPVRKGLAQNFTDITIRFDTGHSNFFLANAQGRGPPIYCSGFCFLTGXREV 60
 QY 60 MPRCTDPLHGPTORRAAQIQAALLGAEREKVEIATFYRKDSCFLGVDPVKNED 119
 D 61 MOKKCSRFILKGPSTSEPALORLHKALEGHOENAEICFRKDSATWCLIDMPINEM 120
 QY 120 GAVIMFLNFEVWEKDWSPADHTNHRGPTSWLAARGAKTRKLPALALTARESS 179
 D 121 GEVVLFLFSFRDIQS---GSPG-----LDP----- 143
 QY 180 VRSGGAGAGAPGAVVVDVLTTPAAPSESLADEVTAMDNHAGLCPAERRALVPGS 239
 D 144 --GGGRDS-----NHNESLG----- 157
 QY 240 PPRAPQQLSPRAHSLNPDASSGCSLARTSRSCQASVRRASADIDEMRAGVLP 299
 D 158 -RRCATWKRFS-----ARRRST-----VL--- 176
 QY 300 PRHASTGAMPLRSGSLNSTSDLVRYRTISKIPQITLNFVLDGDPFLASPTSDREII 359
 D 177 -HRLTG--HGRRG-----GGGMANNVF 198
 QY 360 APKIKERTHNVTEKVOYLSGLADVDEPKIQADPIHRWTILHSPKRAVMDLILVI 419
 D 199 EPK-----PSVEPKVASVSGSRCLLHYSVSKAIDGILLATE 238
 QY 420 YTVNFPYSAFLKETEBEGPAPTEGCAQOPLAIVDLIDIMTIVDILNFRFTYNNAN 479
 D 239 YVAVTVPYNCF--SGDDDPITS-----RHVLVSIAVEMLETLIDILNFRFTYVSQS 290
 QY 480 EEVVSHPRGIAVHYFKGMFLIDWVAIIPFDL-IFGSGSEELGLTARTARLRYVAR 538
 D 291 GQVISARSRISGHLAWFLFDLIALPFDLITFNITVSVLHLLTVRLRLRLLOK 350
 QY 539 LDYSEYGAVALFLKCTPALIAHMLACIWAIGNMOPHMDS---RIGWLHNGDOI 595
 D 351 LERYSCSAVALTLIMSVFALIAHMACIWIIGREMEANDPLMDIGWHEHGRLE 410
 QY 596 PYNSSGGSPSIRKQKYVTALEYFPSSILVSGFGVVSNTNSEKIFESICVMLIGSL 655
 D 411 PYNGSVGSPSRKSAVTAALFTYLSLTSVGFVNCANTDAEKTFTICTMLGALMAV 470
 QY 656 FGNVSAIQRLYSGTARTHTOMLVREFIRHQIPNPLRQRLSEYFOHANSYTGIDMA 715
 D 471 FGNVTALIQMYSRSLYHSMKDLKDFIRHRLPRKLQRMLEYFTTAAVNSGIDANE 530
 QY 716 VLKGFPECLADICLHNRSLQHCPRFGATKCSLALAMKRTTTHAPREDTLVHAG 775
 D 531 LTRFPEDELRAIDIAHMLNREITLQ-LPLFGAASRCLRALSIHKTSCARGEYDLRR 589

```

OY 776 LTALEFISRGSEILIRGCVVVAIIKKNDFG-----EPNLNYPARG--KSNQDVA 824
DB 590 LOAHYVYSGSLVLEIRDMNVLAIIKKGDLIGADIDPEQGP-GLGADPFVLTSTADYKA 648
OY 825 LTYCDLHRIHRDLDLEVDWYEPFSDHFWSSL-ETTFNLRTNMIPGSPGSTELEGGS 882
DB 649 LTYCGLQQLSSNGLAELVRLIPEYGAARRAGLPDITFLRQ-----GSOTS-----GIS 698
OY 883 R-QRRKRLSFRRR---TDKTDQEGEVSALGPGRAGAPSSR-----GRPG-- 925
DB 699 RFSRSPRLSQPSSESLSSSDKTLPSITEAESGAEPGGPRRRPRLPLNLSPARRGSL 758
OY 926 -----PMG-----ESPSSGS-SP-----ESSEDEBPGSSS-----PLALVPPSSSR 962
DB 759 VSLGEEPLPEPSALVSSPSLSPLSPLALAGOGHSASPHPCASAMKPPOLLIP--PL 815
OY 963 PGGEPPGGEPP-LMEDCEKSPDTCNPLSGAFSGVSNFSPWGSRRGROXOELPCAPRPS 1021
DB 816 GTFGPPDLSPRIVDIEDSGTAEPFRFS-----RRPELFRPSQAP- 859
OY 1022 LINTPLSSPGRRRRPGDVSRLDALQRLNRL-----TRLSDMATVLOLLQRCM 1071
DB 860 -----PTGTRPSPELASEAEVEKVCRLNQEISRLNQEVSQSLRELHIMGLQAR- 911
OY 1072 TLVPRAYSAVTPGCGPTSTPLVSPPTITLDSISQVSPMACCELPAPAPLPQSG 1131
DB 912 -LGPPGHPAGSAMPDP-----PCPOLRPPCLG-----PCASRNP--PSL-DPT 951
OY 1132 PTRRLSLPGOLAL 1145
DB 952 TLAEVHCPASVQTM 965

```

RESULT 13

```

ID KCH4_RAT STANDARD: PRT: 1017 AA.
AC 09RT9: 089048:
DT 28-FEB-2003 (rel. 41, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 4 (Ether-a-go-go-
DE like potassium channel 1) (ELK channel 1) (ELK1) (Brain-specific eag-
DE like channel 2) (BEC2).
GN KCN4 OR ELK1 OR ELK3.
OS Rattus norvegicus (rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Brain cortex;
RA MEDLINE=99043952; PubMed=9824707;
RA Engelard B., Neu A., Ludwig J., Roeper J., Pongs O.;
RT "Cloning and functional expression of rat ether-a-go-go-like K+
RT channel genes.";
RL J. Physiol. (Lond) 513:647-654(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=99386988; PubMed=10455180;
RA Miyake A., Mochizuki S., Yokoi H., Kohda M., Furuchi K.;
RT "New ether-a-go-go K+ channel family members localized in human
RT telencephalon.";
RL J. Biol. Chem. 274:25018-25025(1999).
RN [3]
RP TISSUE SPECIFICITY.
RA MEDLINE=21319165; PubMed=11425889;
RA Saganich M.J., Machado E., Rudy B.;
RT "Differential expression of genes encoding subthreshold-operating
RT voltage-gated K+ channels in brain.";
RL J. Neurosci. 21:4609-4624(2001).
CC -I- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
CC channel. Elicits an outward current, but shows no inactivation.

```

```

CC Channel properties may be modulated by cAMP and subunit assembly.
CC -I- SUBUNIT: The potassium channel is probably composed of a homo- or
CC heterotrimeric complex of pore-forming alpha subunits that can
CC associate with modulating beta subunits.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: Highly expressed in adult testis, and in adult
CC and embryonic brain. In adult brain found in piriform cortex,
CC olfactory tubercle, cerebral cortex, hippocampus pyramidal cells
CC and dentate gyrus and basal ganglia of caudate/putamen and
CC accumbens nucleus. Detected at intermediate levels in lung, spinal
CC cord, and pituitary.
CC -I- DEVELOPMENTAL STAGE: Expressed at day E18 in embryonic brain.
CC -I- DOMAIN: The segment S4 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -I- SIMILARITY: Belongs to the potassium channel family. H (dag)
CC subfamily.
CC -I- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -I- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -I- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: AJ007628: CAA07587.1; -.
DR EMBL: AB022699; BAA83593.1; -.
DR PIR: T31354; T31354.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR Pfam: PF00785; PAC; 1.
DR SMART: SM00100; CNMP; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 1.
DR TIGRFAMs: TIGR00229; sensory_box; 1.
DR PROSITE: PS50042; CNMP_BINDING_3; 1.
DR PROSITE: PS50112; PAS; 1.
DR PROSITE: PS50113; PAC; 1.
KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW Potassium channel; Potassium; Potassium transport; Transmembrane;
KW Glycoprotein; Multigene family.
FT DOMAIN 1 232
FT TRANSMEM 233 253
FT TRANSMEM 263 283
FT TRANSMEM 284 305
FT TRANSMEM 306 326
FT TRANSMEM 335 355
FT DOMAIN 356 364
FT TRANSMEM 365 385
FT DOMAIN 428 448
FT TRANSMEM 455 475
FT DOMAIN 476 1017
FT DOMAIN 14 90
FT DOMAIN 93 145
FT NP_BIND 557 672
FT CARBOHYD 415 415
FT CONFLICT 825 825
SQ SEQUENCE 1017 AA; 111403 MW; BFD80F1B35437C9F CRC64;

```

```

Query Match 26.6%; Score 1616; DB 1; Length 1017;
Best Local Similarity 33.6%; Pred. No. 1,9e-76;
Matches 417; Conservative 172; Mismatches 341; Indels 312; Gaps 36;

```


RN [4]
 RP TISSUE SPECIFICITY.
 RX MEDLINE-21319165; PubMed-11425889;
 RA Saganich M.J., Machado E., Rudy B.;
 RT "Differential expression of genes encoding subthreshold-operating
 RT voltage-gated K⁺ channels in brain."
 RL J. Neurosci. 21:4609-4624(2001).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Elicits a slowly activating, outward rectifying current.
 CC Channel properties may be modulated by cAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterooleptimeric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits.
 CC -1- TUBELLULAR LOCATION: Detected in superior cervical, mesenteric and
 CC coeliac ganglia. Expressed in brain (piriform cortex, olfactory
 CC tubercle, cerebral cortex, hippocampus pyramidal cells and
 CC dentate gyrus and basal ganglia of caudate/putamen and accumbens
 CC (nucleus). Expressed in pituitary.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Bag)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF061957; AAC61520.1; -
 DR EMBL; AJ007632; CA07591.1; -
 DR PIR; T17367; T17367.
 DR InterPro: IPR000595; cNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_channel.
 DR InterPro: IPR003280; K+channel_2pore.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS_assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam; PF00520; cNMP_binding.1.
 DR SMART; SM00100; cNMP.1.
 DR SMART; SM00086; PAC.1.
 DR TIGRfam; TIGR00229; sensory_box.1.
 DR PROSITE; PS50042; cNMP_BINDING_3; 1.
 DR PROSITE; PS50113; PAC.1.
 DR PROSITE; PS50112; PAS; FALSE_NEG.
 KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Glycoprotein; Multigene family.
 FT DOMAIN 1 225
 FT TRANSMEM 226 246
 FT TRANSMEM 256 276
 FT DOMAIN 277 298
 FT TRANSMEM 299 319
 FT TRANSMEM 328 348
 FT TRANSMEM 349 353
 FT TRANSMEM 354 374
 FT DOMAIN 420 440
 FT TRANSMEM 449 469
 FT DOMAIN 470 1102
 FT DOMAIN 18 90
 FT DOMAIN 93 145
 FT DOMAIN 711 722
 FT POLY-GLU.

FT NP_BIND 551 668 CNMP.
 FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 71 71 E -> L (IN REF. 2).
 FT CONFLICT 187 187 K -> N (IN REF. 2).
 FT CONFLICT 296 296 I -> T (IN REF. 2).
 FT CONFLICT 370 370 M -> I (IN REF. 2).
 SQ SEQUENCE 1102 AA; 123230 MW; A135CC36E2EF1A3 CRC64;
 Query Match 25.5%; Score 1551; DB 1; Length 1102;
 Best Local Similarity 32.6%; Pred. No. 4.9e-73;
 Matches 383; Conservative 169; Mismatches 283; Indels 340; Gaps 30;
 QY 1 MPVRRGVAPQNFELDTIRKFEQSGKSRFTIARV-ENCVAIYCNDFCELCGYSAEV 59
 DB 1 MPVKGILAPQNFELDTIARFEGTSHNFTLANAQVAKGPPIYCSGFEELAGFATVEV 60
 QY 60 MRRPCDGLHGRPTORRAAQAQALGAEERKEVIAFYRKDSCFLVDVVPVKNED 119
 DB 61 MRSKSCCKFLGVEYTNQMLQLEKSLSEKVEKEGELIMFYKKNAGPWCILDIVPIKNEK 120
 QY 120 GAVIMFLNFEVYMERDMVSGPAHDTNRGPPTSMLAGAKTFRLKPLALITARESS 179
 DB 121 GDVVLPLASF-----KDIPTKY----- 138
 QY 180 VRSGGAGAGAPGAVVVDVDTLPAPSSSLADEVTAMDNVHAGLGPAPERALVPGS 239
 DB 139 -----KITSEK-----KEDR----- 149
 QY 240 PPRAPQQLPPRAHSLNPDAGSSCSGLARTRSRSCASVRRASADDEAMRAGVLRPP 299
 DB 150 ---AKGR-----SRAGSHFDSARRSR----- 168
 QY 300 PRAHSGAMPRLSGLNLSDSDDLRYKRTISKIPQTLNFPVLDKGPFLASPTSDRETI 359
 DB 169 -----AVLYHISGHLDREKKRLKINNVPFVDPK----- 197
 QY 360 APRKERTHNVTEKVTQVLSGADVLEPYEQAPRIHRTIILHSPKAVAMDILLLVI 419
 DB 198 -----APPEYKVSDAKKKFFLLHFSTFKACMDLILATF 233
 QY 420 YTAVFPTYSAAFLFKETEGEPATECGYACQPLAVVDLIVDIMEYDILINFRTYVAN 479
 DB 234 YVAVTPVYVNCIFGNEIDLSTTRST-----TVSDIAVEILFIIDILINFRTYVKS 284
 QY 480 EEVVSHPRIAVYFPKGFLLIDVMAAIPEDLL-IFSGSSEELGLKTAALLRVRARK 538
 DB 285 GGVIFEARSLCIHYVTWTFIIDLIAALPFDLLAFVNTVVSIVHLKTVRLRLRLLOK 344
 QY 539 LDRYSEGAIVLEFLMCTFALLIHWLACIYVATGNMOPH---MDSRIGLHNLGDQIGK 595
 DB 345 LDRYSOHSITVLTLLSMFALLIHWMACIYVIGKMERENDSLKWEVGMHELKGLDES 404
 QY 596 P-YNSGSLGSPSIRKDYVYALYFTFSSLTSGVSGNVSPTNSEKIFSIYVMLIGSLMYAS 654
 DB 405 PYYGNNMTLGPSPRSRAVIALYFTLSLTSGVSGNVSANDAEKIFISICMLIGALMHAL 464
 QY 655 IFGNVSAIITRLYSGARITYTQMLRYREFIRFHOIPNPLQRLLEYFQHMWSTNGIDAN 714
 DB 465 VFGNVNVAIIITORMYSRMSLYTRTKDKDFRIVHNLPOOLKQRMLEVFQITWVSVNGIDSN 524
 QY 715 AVKGPPECQAOADICLNLNSLLOHCKPFGATRGCLRALAMKFKTHAPGDTLVAGD 774
 DB 525 ELAKDPDELRSDTIHLNKEIITQ-LSLFCASRGCRLSLHRTKTSFCAPGEVLQOGD 583
 QY 775 LITALLYFISGSIEILRGDVVAALILKNDIIFGEPLNLYARPGKSNQDVALTYCDLHKIH 834
 DB 584 ALQAIYFVCGSMEVLKDSVLAALILKGDILGANLSTIKDVIVITNADVKALTYCDLQCI 643
 QY 835 RDLLLEVLDMPPEPSDFWSSL--EITFNLRD-----TNMIRPSPGSTELLEG 880
 DB 644 LKGLFEVLYGLYPRYAHKFVEDIQLDLYTLNREGHESDVISRLSNKSTVP---QAEPKGN 699

```

OY      881  FSRORRRRLRRRTDDEOPGVSLGPCRAGAPSSRRPGPMGESPSSGSPSPSS  940
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      700  GS-1KKRLPLSVDEEBEEVEEETTSIP-----IYTRG-----SSVSH  738
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      941  SEDEGPGRSSPLRL-----VPSSP-----RPPGEP-----968
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      739  SKTKGSSKSYGLSLKOLTGCTVPHSHIRVSSANPKTKADPPNHGTRKEKNLYOL  798
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      969  -----GGEPLM-----EDCEKSPDPCNLSAFSGVSNIFSGWMDGRKQOELPRC  1015
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      799  CSLGTACTPELSPRIWDIGEDGNSEET-----OTFPGSE-----QIRPEPRI  842
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      1016  PAPPLSLINPLPSSGRRRGDVE-----SRDLAOROLNLE-----TRISA  1058
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      843  ---SPSL-----GEGEIGAFLFKKEETKQOINKINSEVTLTLOEVSOLGK  886
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      1059  DMATVLOLOROMTLVPAYSAVTPPGGPPSTAP  1093
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      887  DMRSIMOLLENITLSPQPSQPC-----SLHPSTICP  917
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: September 23, 2003, 17:37:46
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 11:05:09 : Search time 88 seconds
(without alignments)
3398.670 Million cell updates/sec

Title: US-10-000-151B-3
Perfect score: 6079
Sequence: 1 MPVRGHVAPQNTFLDITIR.....GOLGALTPSQPLRHGSDPGS 1159

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SPTREMBL_23:*
- 2: sp_archaea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mmc:*
- 9: sp_organelle:*
- 10: sp_phage:*
- 11: sp_plant:*
- 12: sp_protent:*
- 13: sp_virus:*
- 14: sp_vertebrate:*
- 15: sp_unclassified:*
- 16: sp_rv1rus:*
- 17: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5228.5	86.0	1058	4	Q81Z12
2	3241	53.3	1186	13	Q8UH78
3	2430	40.0	732	4	Q81V15
4	1941.5	31.9	791	5	Q18325
5	1865.5	30.7	855	5	002497
6	1530	25.2	515	11	Q8C782
7	1530	25.2	1097	11	Q8BX82
8	1516.5	24.9	522	11	Q8CC38
9	1485.5	24.4	1284	5	Q9V899
10	1485.5	24.4	1284	5	Q23974
11	1453.5	23.9	1174	5	Q9VX26
12	1423	23.4	956	5	Q9YX7
13	1403.5	22.1	988	11	Q8C035
14	1099.5	18.1	395	11	Q8BY52
15	511.5	8.4	1327	5	Q9V702

17	503.5	8.3	678	5	Q96777	Q96777 heliothis v
18	502.5	8.3	945	5	Q9Y1J9	Q9Y1J9 drosophila
19	459	7.6	113	11	Q8BRP8	Q8BRP8 mus musculu
20	440	7.2	767	5	Q76977	Q76977 strongyloce
21	439.5	7.2	900	5	Q97119	Q97119 limulus pol
22	434	7.1	1463	5	Q9U5E2	Q9U5E2 drosophila
23	428.5	7.0	688	10	Q41461	Q41461 solanum tub
24	427	7.0	810	10	Q9XH40	Q9XH40 samanea sam
25	426	7.0	1696	5	Q9YXV8	Q9YXV8 drosophila
26	425	7.0	1324	5	Q8IR35	Q8IR35 drosophila
27	420.5	6.9	611	11	Q9QW7	Q9QW7 rattus norv
28	420.5	6.9	632	11	Q9ER32	Q9ER32 rattus norv
29	420.5	6.9	670	11	Q9ER33	Q9ER33 rattus norv
30	418.5	6.9	631	11	Q8CFV6	Q8CFV6 mus musculu
31	418.5	6.9	1218	5	Q9W201	Q9W201 drosophila
32	412.5	6.8	802	10	Q38898	Q38898 arabidopsis
33	412	6.8	787	10	Q3M0L6	Q3M0L6 arabidopsis
34	409	6.7	686	6	Q9N0H4	Q9N0H4 sus scrofa
35	397	6.5	883	10	Q38998	Q38998 arabidopsis
36	397	6.5	883	10	Q9LE66	Q9LE66 lycopersico
37	393	6.5	883	10	Q24382	Q24382 solanum tub
38	391.5	6.4	677	10	Q42426	Q42426 arabidopsis
39	389.5	6.4	677	10	Q39128	Q39128 arabidopsis
40	389.5	6.4	820	10	Q9FNYS	Q9FNYS arabidopsis
41	387.5	6.4	830	10	Q9FY04	Q9FY04 populus tre
42	385	6.3	849	10	Q9SM12	Q9SM12 zea mays (m
43	385	6.3	887	10	Q04242	Q04242 zea mays (m
44	384.5	6.3	697	10	Q49732	Q49732 arabidopsis
45	384.5	6.3	708	10	Q9C5V9	Q9C5V9 arabidopsis

ALIGNMENTS

RESULT 1	ID	Q81Z12	PRELIMINARY;	PRT; 1058 AA.
AC	Q81Z12;			
DT	01-MAR-2003 (Tremblrel. 23, Created)			
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)			
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)			
DE	Ether-a-go-go related potassium channel.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Colon;			
RA	Shoeb F., Malykhlina A.P., Akbarali H.I.;			
RT	"HERG potassium channel from human colon."			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY130462; RAN05415.1; -			
KW	Ionic channel.			
SQ	SEQUENCE 1058 AA; 116000 MW; 70F37F537AAFE084 CRC64;			
Query Match	86.0%;	Score 5228.5;	DB 4;	Length 1058;
Best Local Similarity	95.1%;	Pred. No. 0;		
Matches 1005;	Conservative	9;	Mismatches 38;	Indels 5;
Gaps				3;
QY	1	MPVRGHVAPQNTFLDITIRKFEQSRKFTIANARVENCAYICNDPFCGLGYSRAEVM 60		
DB	1	MPVRGHVAPQNTFLDITIRKFEQSRKFTIANARVENCAYICNDPFCGLGYSRAEVM 60		
QY	61	QRPCTDHLHGPRTORRAAQAIALGAEERKVEIAFYRKDGSCFLCLVDVVPVKNEDG 120		
DB	61	QRPCTDHLHGPRTORRAAQAIALGAEERKVEIAFYRKDGSCFLCLVDVVPVKNEDG 120		
QY	121	AYIMFVLNEVMEKDMWSSPAHDNRHRRPSTSWLAPGAKFRLLKPLALLTARESSV 180		
DB	121	AYIMFVLNEVMEKDMWSSPAHDNRHRRPSTSWLAPGAKFRLLKPLALLTARESSV 180		
QY	181	RSGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAIGLPABERRALVPG-- 238		


```

181 RTGMSRAGAGAVVADLTPAAPSSSESLADSVSAMDNVAGLGAEEERRALVGGSA 240
239 SPPSAPAGOLPSPRAHSLNPDASGSSCLARTRESRECAVRRASDDIEAMRAGVLP 298
241 SPVASINGPRPSPRAGLNPDSAGSSCPARTKRESCAVRRASDDIEAMRAGALP 300
299 PPRHASTGAMHPLNSGLNSTSDSLVRYRTISKIPOTILNFVDLKGDPLASPTSDREI 358
301 PPRHASTGAMHPLNSGLNSTSDSLVRYRTISKIPOTILNFVDLKGDPLASPTSDREI 360
359 IAPKIKERTHNVEKYQVLSGADVPEYKLOPKIHRWTILHYSPEKAWMDLILLY 418
361 IAPKIKERTHNVEKYQVLSGADVPEYKLOPKIHRWTILHYSPEKAWMDLILLY 420
419 IYAVFPPYSAAFLLKTEEGSPATEEGVACQPLAVVDLYADIMEYDILINFETTYNA 478
421 IYAVFPPYSAAFLLKTEEGSQAPDCGYACQPLAVVDLYADIMEYDILINFETTYNA 480
479 NEEVSHPRGIAVHYFKGWFLLDMVAIPFDLLIFGSGSEELIGLKTARLLRVARK 538
481 NEEVSHPRGIAVHYFKGWFLLDMVAIPFDLLIFGSGSEELIGLKTARLLRVARK 540
539 LDRSEGAVALFLMCTPALIAHWLACIYATIGNMOPHDSKIGLHNGDQIGRPYN 598
541 LDRSEGAVALFLMCTPALIAHWLACIYATIGNMOPHDSKIGLHNGDQIGRPYN 600
599 SSGIGSPISKDYTYTALYFTFSSLTSGFVNSPNTSEKIFSCVMLIGSLMAJSEFN 658
601 SSGIGSPISKDYTYTALYFTFSSLTSGFVNSPNTSEKIFSCVMLIGSLMAJSEFN 660
659 VSAIIOQLYSGTARYHTOMLRVREFIRFHOIPNLRORLEBYEQHANSYTNGLDMNVK 718
661 VSAIIOQLYSGTARYHTOMLRVREFIRFHOIPNLRORLEBYEQHANSYTNGLDMNVK 720
719 GFPECLADICLHNRSLLDHCKFRGATGCLRALAMKRTTTPAPGDTLVHAGDLTA 778
721 GFPECLADICLHNRSLLDHCKFRGATGCLRALAMKRTTTPAPGDTLVHAGDLTA 780
779 LYFISRSIETLRDVAVALIGKNDIGEPNLVAPRGSNGDVRLTYCDLKHINDDL 838
781 LYFISRSIETLRDVAVALIGKNDIGEPNLVAPRGSNGDVRLTYCDLKHINDDL 840
839 LEVLDMPERFSDHFWSSLEITFNLRDNTMIPGSFGSTELGEGFSRQRRKLSERRTKD 898
841 LEVLDMPERFSDHFWSSLEITFNLRDNTMIPGSFGSTELGEGFSRQRRKLSERRTKD 900
899 TEOPGEVSAL--GGRAGAGPSSRGRPGMGSPSSGSPSEDEGPGCRSSPLRV 956
901 TEOPGEVSAL--GGRAGAGPSSRGRPGMGSPSSGSPSEDEGPGCRSSPLRV 960
957 PFSSPPRGPGEPLMEDCEKSSDTCNPLSGAFSVSNIFSWGDSRGROYELPRCP 1016
961 PFSSPPRGPGEPLMEDCEKSSDTCNPLSGAFSVSNIFSWGDSRGROYELPRCP 1019
1017 APTPSLNLPLSSGRRRGDVSRLDALQOLNRL 1053
1020 APTPSLNLPLSSGRRRGDVSRLDALQOLNRL 1056

```

```

Ox NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Langheinfisch u., Wagner T.:
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF532865; AAM95975.1;
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR003967; Erg_channel.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K_channel_pore.
DR InterPro: IPR005820; K_channel_nlg.
DR InterPro: IPR001610; PAC.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR Pfam: PF00785; PAC; 1.
DR PRINTS: PRO1470; ERGCHANNEL.
DR SMART: SM00100; CNMP; 1.
DR PROSITE: PS50042; CNMP_BINDING_3; 1.
DR Ionic channel; Transmembrane.
SQ SEQUENCE 1186 AA; 132279 MW; BCFEF3F23551A8BD3 CRC64;

Query Match 53.3%; Score 3241; DB 13; Length 1186;
Best Local Similarity 57.5%; Pred. No. 1.2e-25;
Matches 717; Conservative 121; Mismatches 252; Indels 158; Gaps 36;

1 MPVRGHVAPQNTFLDTIIRKFEQSRKFIANARVENCAVYICNDGFCELGYSRAEVM 60
1 MPVRGHVAPQNTFLDTIIRKFEQSRKFIANARVENCAVYICNDGFCELGYSRAEIM 60
61 QQSCQCFVPGCGTMSKALQALQALLGSEERKYEIYYSEGCRCCLDVIIVKNEEG 120
121 AVIMEFLNFVYMKDQV--GSPAHDTNNHGRPTSMPLAPGAKFFRLKALLALTRES 178
121 VVIMEFLNFQELDPSKKGGLKORMAN-----SWLRAGORRMHLMPSL--RVKRP 172
179 SVRSAGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVGAGP---AEERRAL 234
173 SLRPHREG-----VVVDV---LQPSHEVALKDLQ-----MSPSCLKSEFQAL 214
235 V--GPGSPRPSAPQQLSPRAHSLNPDASGSSCLARTRESRECAVRRASADDIAMR 292
215 IQQTPSSCELSPPSRPSDRLEPSP-----LTKSHSRESNHSRRASLSLHDIDGM 267
293 AGVLPPEPRHASTGAMHPLNSGLNSTSDSLVRYRTISKIPOTILNFVDLKGDPLASP 352
268 -----DQMSDLKPSNLNSTSDSLMHRITIGRIPQYTIISF-----GSDRLRP 310
353 T-SDREITAP-KIKERTHNYTEK--VTQVLSGADVLPYKLOAPRIHRWTILHYSPEK 407
311 SPTEIETLAVSKIDRSQNVSEKVTQVLSGADVLPYKLOAPRIHRWTILHYSPEK 370
408 AVMWMLLLVLYYAVTTPYSAFLKTEEGSPATEEGVACQPLAVVDLYADIMEYD 467
371 AVMWMLLLVLYYAVTTPYSAFLKTEEGSPATEEGVACQPLAVVDLYADIMEYD 429
468 LINFRTYVYVNAEEVSHPRGIAVHYFKGWFLLDMVAIPFDLLIFGSGSE-----LIG 522
430 LINFRTYVYVNAEEVSHPRGIAVHYFKGWFLLDMVAIPFDLLIFGSGSEDEPOTTLIG 489
523 LKTKARLLRVVYVARKIDRYSEYGAVALFLMCTPALIAHWLACIYATIGNMOPH 579
490 LKTKARLLRVVYVARKIDRYSEYGAVALFLMCTPALIAHWLACIYATIGNMOPH 549
580 -DSRIGWLHNGDQIGRPYN--SSGLGSPSISKDYVYVLYLFPSSLVSVGQANSPNTSE 637
550 GGMKIGWLDNLADQIGROYDNSNFSGSPSISKDYVYVLYLFPSSLTSGFVNSPNTPE 609
638 KIFSICVMLIGSLMAYATIFGVNSAIQRLVSGTARYHTOMLRVREFIRFHOIPPLQRL 697
610 KIFSICVMLIGSLMAYATIFGVNSAIQRLVSGTARYHTOMLRVREFIRFHOIPPLQRL 669

```


RP Locus."
RL J. Neurosci. 17:875-881(1997).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97178302; PubMed=8994043;
RA Wang X.J., Reynolds E.R., Deak P., Hall L.M.;
RT "The seizure locus encodes the Drosophila homolog of the HERG
potassium channel."
RL J. Neurosci. 17:882-890(1997).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Paclab J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Gelinker S.;
RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AEO03462; AAF47148.1; -
DR EMBL: U42204; AAB50956.1; -
DR EMBL: U36925; AAB50936.1; -
DR EMBL: AY058350; AAL13579.1; -
DR FlyBase; FBgn0003353; sel.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR003967; Erg_channel.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K_channel_pore.
DR InterPro: IPR005820; M_channel_nlg.
DR Pfam; PF00027; CNMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR PRINTS; PRO1470; ERGCHANNEL.
DR SMART; SMO0100; CNMP; 1.
DR PROSITE; PSS0042; CNMP_BINDING_3; 1.
KW Ionic channel; Transmembrane.
SQ SEQUENCE 855 AA; 97636 MW; A7251AB786CF3599 CRC64;

Query Match 30.7%; Score 1865.5; DB 5; Length 855;
Best Local Similarity 68.6%; Pred. No. 6.4e-132;
Matches 359; Conservative 64; Mismatches 73; Indels 27; Gaps 7;

QY 355 DRE-TIPKIKERHNATEKYOVLSGADVLPYKIQAPRIHWTLIHSPEFAVMDL 413
DB 270 DKEALLSKSEPKQDNDNMT---SLGNTLIDOKLONNYHKWTLHSPEFAVMDI 326
QY 414 ILLVITYAVTPPSAFLKETEGRPATCEGYA-----COPLAVVDLIYDIMEIV 465
DB 327 ILLVMTALTTPYVAFLGEGD-----IQRRNSKYNSDPYITIDLVDTFIV 377
QY 466 DILINFRTYVNAEEVYVSHPGRIAVHYFKGMFLIDVAAIPFDLLIFGSGSEE---LIG 522
DB 378 DIINFRTYVNSQDEYVSHGRIAVHYLSGMFLIDVAAIPFDLLVGSSTDTETTLIG 437
QY 523 LKTAARLLRYVARAKIDRSEYGAVALFLMCTFALIAHMLACIWAIGMEOPHMSR 582
DB 438 LKTAARLLRYVARAKIDRSEYGAVALFLMCTFALIAHMLACIWAIGMEOSIAKN 497
QY 583 IGMHNLGDOIGRKYNSGSGPSIKDKYVALYFTSSLSVSGFVNSPPTNSEKIFI 642
DB 498 IGMHNLGDOIGRKYNSGSGPSIKDKYVALYFTSSLSVSGFVNSPPTNSEKIFI 557
QY 643 CVMILGSLMVASIFGNVSAIIQRLYSTARYHTOMLVREFIRFHOIIPNPLQRLSEYFQ 702
DB 558 CVMILGSLMVASIFGNVSAIIQRLYSTARYHTOMLVREFIRFHOIIPNPLQRLSEYFQ 617
QY 703 HANSTYNGIDMNAVLKGPPECLQADICLHNRSLQHCPRGATKCLRALAMKEFTTH 762
DB 618 HANSTYNGIDMNAVLKGPPECLQADICLHNRSLQHCPRGATKCLRALAMKEFTTH 677
QY 763 APPGDTLVHAGDILTALYFISRSSEIELR-GDVVVALIGKNDIGEPLNLAYARGKSGND 821
DB 678 APPGDTLVHAGDILTALYFISRSSEIELR-GDVVVALIGKNDIGEPLNLAYARGKSGND 821
QY 822 VRALTYCDLHKIRHDDLEVLDMYPEFSDFHWSLSLETFNLRD 864

DB 736 VRALTYCDLHKIRHDDLEVLDMYPEFSDFHWSLSLETFNLRD 778

RESULT 6
ID 08C782 PRELIMINARY; PRT; 515 AA.
AC 08C782;
DT 01-MAR-2003 (Tremblrel; 23, Created)
DT 01-MAR-2003 (Tremblrel; 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel; 23, Last annotation update)
DE Inferred: Mus musculus potassium channel protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK052366; BAC34958.1; -
SQ SEQUENCE 515 AA; 58246 MW; 113FC2274308E3FE CRC64;

Query Match 25.2%; Score 1530; DB 11; Length 515;
Best Local Similarity 59.4%; Pred. No. 6.6e-107;
Matches 315; Conservative 64; Mismatches 123; Indels 28; Gaps 9;

QY 1 MPVRGHVAPONTFIDITIRKEGOSRFTIARNRVNCAYVNGDPCCLGYSRAEYM 60
DB 1 MPVRGHVAPONTFIDITIRKEGOSRFTIARNRVNCAYVNGDPCCLGYSRAEYM 60
QY 61 QRPCTCFELHGPRTQRRAAQIAQALLGAERKEVIAFYRKDSCFCLYDVVYVKNEDG 120
DB 61 QRPCTCFELHGPRTQRRAAQIAQALLGAERKEVIAFYRKDSCFCLYDVVYVKNEDG 120
QY 121 AVIMEILNFEVYMEKDWGSPADTNHGRPTSWLADQRAKTFPLKLPALIALTARSSV 180
DB 121 VAMFEILNFEVYMEKDWGSPADTNHGRPTSWLADQRAKTFPLKLPALIALTARSSV 175
QY 181 RSGGAGAGAPGAVVVDLTPRAPSSESLALDEVTAAMDNVAGLGAERERALVGG-- 238
DB 176 PQED-----PDVVVID---SSKHSDSVAMKHFKSPYKESCPSEADDTKALIDPSQC 225
QY 239 SPPRSAPGOL---PSPRAHSLNPDAGSSCSLARTSRSCASVRRASSADDIENRAG 294
DB 226 SPLVNIISGPDHSSPKQMDRLYPDMLQSSQLTHSRRESLCSIRRASSVHDIIEG--S 283
QY 295 VLRP---PBRHASTGAMHPLKSGLLNSTSDSLVRYRTISKIPQITLNFYDLKGDPLAS 351
DB 284 VHRKNIIPRDRHASGEPNHIKSSLSIGTSDSNLKNKYSTINKIPQITLNFSDVTKKKNTS 343
QY 352 -PTSDRELLIAPKIKERHNATEKYOVLSGADVLPYKIQAPRIHWTLIHSPEFAV 410
DB 344 PPSSDKITLAPKIKERHNATEKYOVLSGADVLPYKIQAPRIHWTLIHSPEFAV 403
QY 411 DMLILLVITYAVTPPSAFLKETEGRPATCEGYACOPPLAVAVDIDVIMFIVDLIN 470
DB 404 DMLILLVITYAVTPPSAFLKETEGRPATCEGYACOPPLAVAVDIDVIMFIVDLIN 462
QY 471 FRTTYVNAEEVYVSHPGRIAVHYFKGMFLIDVAAIPFDLLIFGSGSEEL 520
DB 463 FRTTYVNAEEVYVSHPGRIAVHYFKGMFLIDVAAIPFDLLIFGSGSEEL 512
RESULT 7
ID 08BX82 PRELIMINARY; PRT; 1097 AA.
AC 08BX82;

```

DF 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE ELK channel 3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
DR EMBL; AK048629; BAC33401.1; -.
FT NON_TER 1
SQ SEQUENCE 1097 AA; 122704 MW; D7E37630C1E93B5F CRC64;

Query Match 25.2%; Score 1530; DB 11; Length 1097;
Best Local Similarity 32.0%; Pred. No. 2,1e-106;
Matches 365; Conservative 163; Mismatches 322; Indels 290; Gaps 18;

QY 6 GHVAPQNTFLDTIIRKFGSGSRKFIITANARV-ENCAYIYQNDGFCGLGYSRAEYVQRPC 64
DB 1 GLLAQNTFLDTIATRFQDTHSNFIITANAVAKGFPYIYCSDFGLAFARTEVYVQKSC 60
QY 65 TQDFLHGPTQRRAAQIQAALLGAERKEVIAFRKQSCFLCYVDVYVKNEDGAVIM 124
DB 61 SCKFLPGVETNEQMLQMLQIKESLEEKVEFEKGTMPFKKNAPWPCLDIYPIKNEGDVVL 120
QY 125 FLTFEEVWEKMDVSGSPADHTNHRGPTSWLAPGAKTFRLKLPALLATARESSVRSOG 184
DB 121 FLASRK----- 126
QY 185 AGGAAPGAVVVDLTPAAPSESLADEVTAMDNNHVAAGLPABERRALVGPSPPSRA 244
DB 127 -----DIDDTYKVKITSEDKKEBRTGRSR----- 150
QY 245 PQLSPRAHSLNPASGSSCLARTRSRESCASVRRASSADDIEMRAGVLP PPPRHS 304
DB 151 -----AGSHEDSARRSR----- 163
QY 305 TGAMHPLBSGLNSTSDSLVRYRTISKIPQITLNFVDKGPFLASPTSDREIITAPKIK 364
DB 164 -----AVLYHISGHLQREKKKIKINNVPYDKP----- 192
QY 365 EETHNVTEKVTQVLSIGADVLPDEYKIQAPRIHRWTLIHSPPKAVWDMLILLVITYAVF 424
DB 193 -----APEYKASDAKSKSFILHSTPKAGWMLILLATFYVAVT 233
QY 425 TPYSAFLIKETEDEGPATEGCGYACQPLAVVDLIYDIMEIVOLINFRITYYNADEVVS 484
DB 234 VYVNCFFIGNEDLSTTRST-----TVSDIAVEILFIIDILINFRITYYKSGQVIF 284
QY 485 HPGRIAVHYFKGMFLIDWVAIPDLL-IFGSGSEELIGLTKTARLLRVRARKDKRS 543
DB 285 EAKRSICIHVYTWTFIIDLAIPFDLIAFNVTYVSLVLLTKVRLIRLRILQKIDRS 344
QY 544 EYGAADVLELCTEALLAHMLACIYVAGNMEOPH--MDSRIGLNLHNGDQIGRP-YNS 599
DB 345 QHSTIVLTLLMSMFLAHMAMACIYIIGKMEREDNSLLKWEVGMHELGKRLSPYGN 404
QY 600 SGLGSPSTKDKVYVLTALYFEFSLTVGCGNVSPNTNSEKIFISICWMLIGSLMYAIFGVY 659
DB 405 NTLGGPSTRSAIYALVFTLSLTIVGCGNVSAINDAEKIFISICWMLIGALHMLALVFGVY 464
QY 660 SAIIORLYSGTARYTOMLRVREFIRFOIIPNPLQRLEVEYFOHMASTYNGIDNNAVLKG 719
DB 465 TAIIDORMSRMSLHYTRKDKLDFIRVHHLPOOLQKRMLEYFOTITWSVANGIDSNELKD 524

```

```

QY 720 PPECIQAIDICHLNLSLQHCPRFGATRGCLRALAMKFKTHAPGDTLVAGDILTRAL 779
DB 525 FPDELRSDITMHLDEKILQ-LSLFECAASRGCRSLSLHKTYSFCAPAGEVILKQGDALQAI 583
QY 780 FYTSGSTIEILGADVVALITGKNDIFGEELNLNYPARGKSGGVRLATYCDLKHIRDDL 839
DB 584 YVCGSGMEVLDKDSMVALITGKDLGKDLGKDLNLSIKDQYIKTNADVKALTYCDLCKILKLF 643
QY 840 EYLDVYPERSDHFWSSL--EITFNLRD-----TNMIPGSPGSTELEGFSRORRR 887
DB 644 EYLGAYPEYAHKFVEDIQHDLTYNLRGHESDVYSRLSNKSTVSGAEPKNGSIN--KR 700
QY 888 KLSFRKRDQKDTQGEVSAALPGRAGAGPSRRGPG----- 925
DB 701 IPSIYDEDEEEVEEETTSLSPIYRGSSVSHSKTKGSKNTYGLSLKQLASGTVPPHS 760
QY 926 PWGESPPSGSPSESEDEGPRSSPLLYVFSSPRPGEPGGEPLMEDECKSDPCN 985
DB 761 PIRVSANSPPKQKQADPNHGRKKKKNLKVOLSSLSGAGTPELSPRIVDGE----- 813
QY 986 PLSGAFSGVSNIFSEWGDGRGQYDELPCRPAPFSLNINPLSSGRRRPGDVEERLDA- 1044
DB 814 -----DGNMNETQTFDFGSEQIRPEPRISPL-----ASEEIGAA 849
QY 1045 -----LQRLNLE-----TRLSADMATVQLQRLQMTLVPPAYSAAVTPGP 1086
DB 850 FLFIRAETKQOINKINSEVTLTQEVSLQGHDSIMQDLLENILSPQPSQFCSLHPTP 909
QY RESULT 8
DB ID Q8CC38 PRELIMINARY; PRT; 522 AA.
AC Q8CC38;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Inferred: Mus musculus potassium channel protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
DR EMBL; AK034003; BAC28543.1; -.
SQ SEQUENCE 522 AA; 59030 MW; D687719147B84D38 CRC64;

Query Match 24.9%; Score 1516.5; DB 11; Length 522;
Best Local Similarity 58.7%; Pred. No. 7,1e-106;
Matches 315; Conservative 64; Mismatches 123; Indels 35; Gaps 10;

QY 1 MPVRGHAAPQNTFLDTIIRKFGSGSRKFIITANARVENCAYIYQNDGFCGLGYSRAEYV 60
DB 1 MPVRGHAAPQNTFLDTIIRKFGSGSRKFIITANARVQNCAYIYQNDGFCGEMTGFSPRPVY 60
QY 61 QRPCTCDLHGPTQRRAAQIQAALLGAERKEVIAFRKQSCFLCYVDVYVKNEDG 120
DB 61 QRPCTCDLHGPTQRRKHDIAQAALLSEERKVEVYTHYHNGSFTICWTHIIPVKNQEG 120
QY 121 AYIMFLTFEEVWEKMDVSGSPADHTNHRGPTSWLAPGAKTFRLKLPALLATARESSV 180
DB 121 VAMMFILTFEEVWEKMDVSGSPADHTNHRGPTSWLAPGAKTFRLKLPALLATARESSV 175
QY 181 RSGAGAGAPGAVVVDLTPAAPSESLADEVTAMDNNHVAAGLPABERRALVGPSP- 238
DB 176 PQED-----PDVYVID-----SKHSDDSVAMKHKFKPKESCSFSEADDTKALIQPSOC 225

```

QY	239	PPRPARPOL----	PSPRASHLNPDAAGSSCSCLARTRRESCASVYRASSADDIEAMRAG	254
Db	226	SPLVNISGPLDHSPPKROWDRILYPMLOJSSSOLTJHRSRSESCISIRASSVHDIIEP--S	283	
QY	295	VL.P-.-P.PRAAS-----TGAMPLRSLGSLNSTSDSLVRYRTISKIPQULINFDLK	344	
Db	284	VHPKNIIFDRHASEDNGRNVKCPFNHIKSSLLGSTDSDMLNKSTINKIPQULINFDLVK	343	
QY	345	GDPEFLAS-PTSDREIIAPKIKERTHNHTEKVTQVULSLGADVLP.EYKLOAPRIHRMTILH	403	
Db	344	TEKNTSPSSDSKTIAPKVKERTHNHTEKVTQVULSLGADVLP.EYKLOAPRIKFWTILH	403	
QY	404	SPFKAVMOMLILLYITAVFPYSAFLIKETEEBPPATEGACAPLAVVDLILYDIME	463	
Db	404	SPFKAVMOMLILLYITAVFPYSAFLILNDRE- QKRECEGYSOSP.LAVVDLILYDIME	462	
QY	464	IVDILINFRITYVANEEVYSHPRITAVHYFKGWLIDMVAALIPDULIFGSGSEEL	520	
Db	463	IIDILINFRITYVNONEEVSDPAITAIHYFKGWLIDMVAALIPDULIFGSGSSDEV	519	

RESULT 9

ID	Q9Y899	PRELIMINARY;	PRT:	1284	AA.
AC	Q9Y899;				
DT	01-MAY-2000 (TReMBLrel. 13, Created)				
DT	01-OCT-2002 (TReMBLrel. 22, Last sequence update)				
DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)				
DE	CG5076 protein.				
GN	ELK OR CG5076.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-Berkeley:				
RX	MEDLINE=20196006; PubMed=10731137;				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Peiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abrial J.F., Agbayandi A., An H.-U., Andrews-Piannoch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brooksline P., Brotter P.,				
RA	Burtis K.C., Busam D.A., Butler H., Catleu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.				
RA	Dubin K.J., Evangelista C.C., Ferriz C., Ferreira S., Fleischmann W.				
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,				
RA	Jatalai M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Laško P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,				
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,				
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,				
RA	Spyrakis R., Spradling A.C., Stemple M., Strong R., Sun E.,				
RA	Stevens R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,				
RA	Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissenbach J.,				
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,				
RA	Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,				

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.?
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).

RN [2]
 RP SEQUENCE FROM N.A.
 RA Amanatides P.G., Brandon R.C., Rogers Y., An H., Baldwin D.,
 RA Barzon J., Beeson K.Y., Busam D.A., Carlson J.W., Center A.,
 RA Champe M., Davenport L.B., Dietz S.M., Dodson K., Dorsett V.,
 RA Doup L.E., Doyle C., Dresnek D., Farfan D., Ferreira S., Frise E.,
 RA Galle R.F., Gary N.S., George R.A., Gonzalez M., Houck J.,
 RA Hoskins R.A., Hostin D., Howland T.J., Ibegwam C., Jatali M.,
 RA Kruse D., Li P., Mattei B., Mostrefi A., McIntosh T.C., Moy M.,
 RA Murphy S., Nelson C., Nelson K.A., Nunoo J., Paolel J., Parasas V.,
 RA Park S., Patel S., Pfeiffer B., Phouanavong S., Pittman G.S.,
 RA Puri V., Richards S., Scheeler F., Stapleton M., Strong R.,
 RA Svirskas R., Tector C., Tyler D., Williams S.M., Zaveri J.S.,
 RA Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.
RA Mira S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.
RA Hradecky P., Huang Y., Kaminker J.P., Prochnik S.E., Smith C.D.,
RA Tury J.L., Bergman C.M., Bennett B.J., Carlson J.W., Celnik S.E.,
RA Clamp M.E., Dysdale R.A., Emmert D., Fitse E., de Grey A.D.N.J.
RA Harris N.L., Krommiller B., Marshall B., Milburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smutnack F.,
RA Whitefield E.J., Yamada C., Ashbourne M., Gelbart W.M., Rubin G.M.,
RA Mungall C.J., Lewis S.E.
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.

DR EMBL; AE003801; AA57772.2; -.
DR FlyBase; FBgn0011589; elx.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR003667; Erg_channel.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR003280; K+channel_2pore.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR005820; M+channel_nlg.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000700; PAS-assoc.C.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00027; cNMP_binding; 1.
DR Pfam; PF00520; ion_trans; 1.
DR Pfam; PF00785; PAC; 1.
DR PRINTS; PR01333; 2PORECHANNEL.
DR PRINTS; PR01470; ERGCHANNEL.
DR SMART; SMO0100; cNMP; 1.
DR SMART; SMO0086; PAC; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS50042; cNMP_BINDING_3; 1.
DR PROSITE; PS50113; PAC; 1.
DR PROSITE; PS50112; PAS; 1.
DR Ionic channel; Transmembrane.
SO SEQUENCE 1284 AA; 141378 MW; C6EC31891E4B6D24 CRC64;

Query Match	24.48;	Score 1485.5;	DB 5;	Length 1284;
Best Local Similarity	32.38;	Pred. No. 6.2e-103;		
Matches 366; Conservative	173;	Mismatches 330;		
		Indels 265;	Gaps 27	

[illegible]

[illegible]

01-MAY-2000 (TREMBLrel. 13, Created)
01-MAY-2000 (TREMBLrel. 13, Last sequence update)
01-OCT-2002 (TREMBLrel. 22, Last annotation update)
EAG protein.
EAG OR CG10952.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abdl J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,
RA Bokorova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Ciesley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doul L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudkin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Flocker C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glaser A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Kapran G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003497; AAF8410.1; -.
DR FlyBase; FBgn0000535; eag.
DR InterPro; IPR0000595; cNMP_binding.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR001622; K_channel_pore.
DR InterPro; IPR005820; M_channel_olig.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000700; PAS_assoc.C.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF000027; cNMP_binding.1.
DR Pfam; PF00520; Ion_trans.1.
DR Pfam; PF00785; PAC.1.
DR SMART; SM00100; cNMP.1.
DR SMART; SM00086; PAC.1.
DR SMART; SM00091; PAS.1.
DR PROSITE; PSS0042; cNMP_BINDING_3.1.
DR PROSITE; PSS0113; PAC.1.
DR PROSITE; PSS0112; PAS.1.
DR Ionic channel; Transmembrane.
SQ SEQUENCE 1174 AA; 126370 MW; 88f86ACBB5627FD3 CRC64;


```

Db 121 DAVVLYLCOFK----- 131
OY 180 VRSGAGAGAGAVVVDVLTLPAPSSSESLALDEVTAANDHNHAGLGAERRALVPGS 239
Db 132 -----DITP----- 135
OY 240 PPRAPGOLPSBRAHSLNPDASGSSCSLARTRESASCASVRRASSADDIEMRAGVLP 299
Db 136 -----LQOP----- 139
OY 300 PRHASTGAMHPLRSLNSTSDSLVRYRTISKIPQITLNVYDLKGDPPLASPTDREIT 359
Db 140 -----LDDENKGLSRILQIARIKSKQOENQITKDLHKSP----- 176
OY 360 APKIKERTHNVEKTOYLSIGADVLPKYLOAPRIHRTILHSPFKAVMWMLLLVI 419
Db 177 -----GNTSSNFNOVMNLGDMLPQROETPTSPHILHYSSEFTIMWSTLALTE 228
OY 420 YTAVFTPYSA-----FLKETECPATECGACOPLAVAVDILVIMFTVILINFR 472
Db 229 YTAEVFPENIAFKNSLRFPYLLSSREN-----GGIDSVALMDSIVDVIFFADILLNFH 283
OY 473 TTYVANAEVVSHPGRIAVHYFKWFLIDMVAIPFDLLIEGSGSEELIG-----LKTAR 528
Db 284 TTFVGPGEVVEIPEFVIRQNFYKSWFLIDLISCLPYDIFYMKRDERDISLFSALKVVR 343
OY 529 LLRLVRAVRKIDRYSEYGAVALFLMCTFALAHMLACIVYAGME-QPHNDSKI---G 584
Db 344 LLRLGRVARKLDNYLEYGAAATLLLLCAVYAHMLACVWFIQSEVRLKMDNALPDG 403
OY 585 WLNHNGDQIGKPYN-----SSGLGCPSTIKDKYVVALYFTFESLSLVGSGNVPPTNSEK 638
Db 404 WLMKLSNDLRQVNYNPLSKKTTLVGGPSTSAVISLVTYMGCMVTGCGNASTMTDNK 463
OY 639 IFTSVMLIGSLMTASTIGNVSATIQRLXSGTARYHTOMLKYREFIRFHQIPNPLRQLE 698
Db 464 IFGVOMITISALLYALFQHMHTIIQOMTSSIVRYHEMISNVRFEIKLOEIPKEILAEVYM 523
OY 699 EYFOHAWYTNGLIDMNAVLKGFPECLQDICHLRSLLOHCKPRGATKGLRLAMKF 758
Db 524 DTVSTWMTKIDYAKVGYCPKMKADICVHLKRNKFNHESCRSLSDGCLRLAMFL 583
OY 759 KTTNAPGDTLVHAGDILTALYFISGTEILRGDVVAALIKNDIFGEPL-NLYARPGK 817
Db 584 ELNHAAPQDLYHGESVDALMFVYSGSLVQDDEVAAILKGQVGFGEFKANGSTGQ 643
OY 818 SNGDVRLATYCDLHKIHRDILLEVDMPESSDHFWSSLEITFNLRTNMIPSGSTEL 877
Db 644 SAANVRALTYSDLHMIKKDKLMDVDLFYKAFANSFARNMTLTYNLTH-RMKFRVADYKR 702
OY 878 ECGFSRQK-----RKLSFRR-----TDKTEQGEVYALGP 910
Db 703 EELDAKKRNEKLLPNHPIRKLLEFRMERHGPRIFFSPMADIEKGKTKTETLSRI-- 760
OY 911 GRAGAPSSRRGPGMGESPSSGSPSSSEDECPGSSPLRLVFPSSPGPEPGCG 970
Db 761 -----SLSHSMIDETGGGSSVYK-SPKSKPRP----- 788
OY 971 EPLMEDCEKSDPTCNPLSGAFSGVSNIFSFWG-DSRGROYOLPRCPAPFTPLNIPSS 1029
Db 789 -PLMKQTVDEDLARTS-----WGMDDKREWSLSNKTETMKSKFDI----- 831
OY 1030 PGRRRP--GDVESRLDALOR-----QNLRLERPLSDMATVYOL 1066
Db 832 IGERLTIIEQINSRLALLERVLIGNNGANPSTMPVGSFALSNGNLTLDAAPVARS 891
OY 1067 L-----OROMTLVPPAYS-----AVTTGPGPTSTPLP 1096
Db 892 VMSQOHQPHMQRTST-VPLRLRELAGEMERPIRREPTNPSTSSSRVP 938

```

RESULT 13
044164

```

ID 044164 PRELIMINARY; PRT; 956 AA.
AC 044164;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 108.1 kDa protein.
GN F16B3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium";
RL Science 282:2012-2018(1998).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Kemp K., Wilson R.;
RT "The sequence of C. elegans cosmid F16B3.";
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL AF036695; AAB8348.2; --
DR WormPep: F16B3.1; CE24905.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc.C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00027; CNMP_binding_1.
DR Pfam: PF00785; Ion_trans_1.
DR SMART; SM00100; CNMP_1.
DR SMART; SM00086; PAC_1.
DR SMART; SM00091; PAS_1.
DR TIGRFAMs; TIGR00229; sensory_box_1.
DR PROSITE; PS00042; CNMP_BINDING_3; 1.
DR PROSITE; PS00113; PAC_1.
DR PROSITE; PS00112; PAS_1.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR Hypothetical protein; Ionic channel; Transmembrane.
SQ SEQUENCE 956 AA; 108090 MW; E998B7F2128054A8 CRC64;

```

Query Match 23.4%; Score 1423; DB 5; Length 956;
Best Local Similarity 30.1%; Pred. No. 2,1e-98;
Matches 357; Conservative 166; Mismatches 323; Indels 342; Gaps 26;

```

OY 1 MPV-RGHVAPONTFLDTIRFEQOSRKFIITANARVENCAYIYNDGFCELGYSRAEV 59
Db 1 MPVGRGLVAPONTFLENVIRRCNNADTSFILANQVVDYPIYVNCNDSKLVGYTRABI 60
OY 60 MQRCTCQFLGPRFKYQRRAAQIAQALLGAERKYEIAFYRKDSCFLCVVYVYKND 119
Db 61 MOKPCSLAFMGEHEVEVSLOKMOBALENARTEQAEIGLCKKNKPIMLVLAFLIKNHK 120
OY 120 GAVIMEIINFVEVMEKDWGSPAHDTNHRGPTSM LAPRAKTFRLKLPALLALTRARESS 179
Db 121 DAVVLYLCOFK----- 131
OY 180 VRSGAGAGAGAVVVDVLTLPAPSSSESLALDEVTAANDHNHAGLGAERRALVPGS 239

```

Db 132 -----DITP----- 135
OY 240 PPSAPQOLSPRAHSLNPDASGSSCSLARTSRSCASVARRASSADDIEAMRAGVLP 299
136 -----LKP----- 139
OY 300 PRAHSTGAMHPLNSGLNSTSDLVYRTISKIPQITLNFVDIKGDFLASPTSDRELI 359
140 -----LDDENNKLSRILOIARIASKOOFNOJETDKLHSP----- 176
OY 360 APIKERTHNVTEKVOVLSIGADVLEPKLOAPRIHRTIHAHSPKAWMDLILLLVI 419
177 -----GNTSSNFQVMNLGGMLPQYQOETPKPSPHILHSSKFTMDWSILATP 228
OY 420 YTAFTPYSA-----FLIKETEGRPATCEGYACOPLAVDLIDIMEYIILNFR 472
229 YTAFTVFNIAFKNSLRPFILISRENP-----CGGIDVALMDSIYDVFADILNFI 283
OY 473 TTYVANEVYSHBGRJAVHYFKGFLIDVNAIPDILIFGSGSEELIC---LKTAR 528
284 TTFVPGGEVIEPSVIRONYFKSMFLIDLSCLPYDIFYWFKRDDRIGSLFSLKVR 343
OY 529 LIRIVARAKLDRSEGAVALFLMCTFALIAHMACIYAGNME-QPHMDSRI---G 584
344 LIRIGVARKLDNLEYGAATLLELCAVYIVAHMACVWFWDIDSEVRLKMDLALPDG 403
OY 585 WLNHAGDOIGKPYN-----SSGIGSPSIDKRYTALYFTFSSLTSGFQNVSPNTSEK 638
404 WMLKSLNDLRQHYIPLSNKTTLYGSPSRISAYISLITMCSMTSTYGFGLIANTNEK 463
OY 639 IFSICVMLIGSLMAYISFGVNSAIIQRLYSGTARYHOMLRVREFIRHOIPLNPKOLE 698
464 IFGCMAMISALTAALFGHMTTIIQOMTSYVRYHEMISVREFIKOETPKLAERV 523
OY 699 EYFHANSYNGIDMNVLAGFPECLQADICLHLNRSILOCKPFRGATKCALMAKF 758
524 DYVSTYAMTKGIDITAVLGCPCDKADICVHLNRKVFENHSCFRLASDGLSLAMEL 583
OY 759 KTHHAPGDTLVNAGDILLTALYFISRSIELRGDVVAALGKNDIGEPL-NLYARPG 817
584 ELNHAAPGDLHYHTGESVDALMFVSGSLEYIODDEVVAALGKGVDFEWMKANGSTQ 643
OY 818 SNGDVRLTYCDLAKIRDDILEVDMYPERSDHFSLSLETFFLROTMMIPGSPSTEL 877
644 SAANVRALTYSDLMIMIKKLDLVLYKAFANSFARNMUTLYNLTH-RMFRKRVADYK 702
OY 878 ECGFSRPR-----RKLSFR-----TDKTOEGEVSALGP 910
703 EKELDARKNEKLTLPNDHPIRLKLFMRERHGRIPSPMFADIEGKSTISRI-- 760
OY 911 GRAGAPSSRGRPGPGWGESPPSSPESSEDEGPRSSPLRLVFPSSPRPGEPPGG 970
761 -----SSLHMDIETGGGSSYK-SPRSKKRP----- 788
OY 971 EPLMEDEKSSDTCNPLSGAFSGVSNIFSPWG-DSRGROYOLEPRCAPTPSLNLPISS 1029
789 -PLMKROTVEDDALSRYS-----WGMKDKREWSLSLNIKTEMKSKPDI---- 831
OY 1030 PGRMR--GDVESHDLALOR-----QLRLERLSADATYLOL 1066
832 IGERLTIIEOINSLALIERVLIGNNGANTPSTMPVGSFALSSEGNRLADAPARS 891
OY 1067 L-----QROMTLVPPAYS-----AVTTPGPGPTSTPLIP 1096
892 VSMSEQHOPHKQRTST-VPLRLRELAGEWEPRIREPFPNPTSSSRYP 938

RESULT 14
ID 08C035 PRELIMINARY; PRT; 988 AA.
AC 08C035;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE EAG K+ channel.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RX MEDLINE=2254683; Pubmed=12466851;
RA THE FANTOM Consortium.
RA the Riken Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL: AK032438; BAC27869.1; "
SO SEQUENCE. 988 AA; 11781 MW; 0EEB1F57F5C18BB CRC64;

Query Match 23.18; Score 1403.5; DB 11; Length 988;
Best local similarity 28.88; Pred. No. 6.5e-97;
Matches 354; Conservative 177; Mismatches 337; Indels 363; Gaps 24;

OY 4 RGHVAPONFELDTIIRKFEQSKFTIANARVENCAYICNGFCGLGYRAEVQRP 63
Db 5 KRGLVAPONTFLEIVIRSRSESS--FLLGNAQIDVMPVVISNDFCKLSGTHRADVMOKS 62
OY 64 CTCDFELHPTORAAQIALGAEERKEVETAFYRKDSCFCGLDVVPVKNEDGAVI 123
Db 63 STCSFMGELDLDKTKIEKVAQTFDNYESNCEVLLYKKNTPWYFQIAPIRNEHKV 122
OY 124 MFLINEVYMEKDNVSPAHDTNHRGPTSWLALGAKTRTLKLPALLALTARESSVRSG 183
Db 123 LFLCTFEDI--LFLKOPIEDDSTYK--WTKRAR----- 151
OY 184 GAGAGAPGAVVVDVDTTPAPSESESLDEVTAMDNVHVAIGLPAEERALVGGSPPRS 243
Db 152 -----LTRLATNSRV-LOQLPDM----- 170
OY 244 APQOLSPRAHSLNPDASGSSCSLARTSRSCASVARRASSADDIEAMRAGVLP 303
Db 171 ----- 170
OY 304 STGAMHPLRGLNSTSDSLVRYRTISKIPQITLNFVDIKGDFLASPTSDRELIAPKI 363
Db 171 ----- 170
OY 364 KERTHNVTEKVOVLSIGADVLEPKLOAPRIHRTIHAHSPKAWMDLILLLVIYAV 423
Db 171 KETVYHKSRLAEVLOGLSDILPQYKQDAKTRPHIILHCARFTTMDWYILITFTYAI 230
OY 424 FTYPSAFLKETEEGRPATECGYACOPLAVDLIDIMEYIILNFRTTYVANEV 483
Db 231 MVRVNVSAFKONNIA-----WVLDSVVDVIFLVDIVLNFHTFVPGGEV 278
OY 484 SHGRIVAHYFKGFLIDVNAIPDILIFGSGSEELIGLTKARLLRLVRAVAKI 539
Db 279 SDPLRIMNYLKTFFVIDLSCLPYDIINAFENVDEGISLFSLLKAVRLRLGRVARKL 338
OY 540 DRYSEGAVALFLMCTFALIAHMACIYAGNME-----QPHMDSRIQLHNDQ 592
Db 339 DHTLEIGAVAVLVLYCVGLVAHMLACIWSIGDYEVIDEVNTIQIDS--WLYOLALS 395
OY 593 IGKPY--NSGSL--GGPSIKDKVYVLTALYFTFSSLTSGFQNVSPNTNSEKISICYMLI 647
Db 396 IGTPYRYNTSAGIIEGSPSDSLYVSSLTFTMNSLITIGGNALPPTDVKMPSVAMMY 455
OY 648 GSLMTASIFGVNSAIIQRLYSGTARYHOMLRVREFIRHOIPLNPKOLEEYFOAHV 707
Db 456 GSILYATIFGVNTTIFQOMYANTNRYHEMNNVADFUKLYOVPGLSERWADYIVSTWS 515
OY 708 TNGIDMNAVLYKGPECLQADICLHLNRSILOCKPFRGATKGCRLALMEKTTNAPGD 767
Db 516 SKGIDTEKVLISICKPKRRADICVHLNRKVFENHAPFOLASDGLRALAVEPOTIHCAPGD 575

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 17:35:40 ; Search time 37 Seconds

(without alignments)
1325.359 Million cell updates/sec

Title: US-10-000-151b-3

Perfect score: 6079
Sequence: 1 MPVRGHVAPQNTFLDTIIR.....GOLGALTSPDHRHSGSDPGS 1159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6079	100.0	1159	2	US-08-956-242-13
2	6079	100.0	1159	3	US-09-351-215-13
3	6079	100.0	1159	3	US-09-226-012-2
4	6079	100.0	1159	3	US-09-226-012-4
5	6079	100.0	1159	4	US-09-358-383C-10
6	2436.5	40.1	888	2	US-08-956-242-4
7	2436.5	40.1	888	3	US-09-351-215-4
8	1653	27.2	626	2	US-08-956-242-2
9	1653	27.2	626	3	US-09-351-215-2
10	1636	26.9	1017	4	US-09-600-776-6
11	1568	25.8	1107	4	US-09-358-383C-16
12	1551	25.5	1102	4	US-09-358-383C-36
13	1541.5	25.4	1083	4	US-09-343-494-1
14	1540.5	25.3	1083	4	US-09-600-776-2
15	1540.5	25.3	1083	4	US-09-358-383C-2
16	1536	25.3	1082	4	US-09-336-643A-20
17	1485.5	24.4	1284	4	US-09-343-494-9
18	1485.5	24.4	1284	4	US-09-358-383C-11
19	1415	23.3	962	4	US-09-614-480-9
20	1399.5	23.0	988	4	US-09-614-480-2
21	1190	19.6	542	4	US-09-358-383C-5
22	716.5	11.8	241	4	US-09-358-383C-23
23	565	9.3	181	4	US-09-358-383C-26
24	520	8.6	749	4	US-08-997-685A-10
25	515.5	8.5	910	4	US-08-997-685A-2
26	470	7.7	528	4	US-08-997-685A-4
27	460.5	7.6	290	4	US-09-358-383C-8

28	460	7.6	597	4	US-08-997-685A-12	Sequence 12, Appl
29	445	7.3	506	4	US-08-997-685A-6	Sequence 6, Appl1
30	436.5	7.2	319	4	US-09-358-383C-22	Sequence 22, Appl
31	351.5	5.8	180	4	US-09-358-383C-24	Sequence 24, Appl
32	334	5.5	189	4	US-09-358-383C-25	Sequence 25, Appl
33	315.5	5.2	126	4	US-08-997-685A-59	Sequence 59, Appl
34	303.5	5.0	58	4	US-09-358-383C-29	Sequence 29, Appl
35	248	4.1	52	4	US-09-358-383C-35	Sequence 35, Appl
36	233.5	3.8	89	4	US-09-358-383C-30	Sequence 30, Appl
37	232.5	3.8	170	4	US-09-358-383C-27	Sequence 27, Appl
38	216	3.6	75	4	US-09-358-383C-28	Sequence 28, Appl
39	214	3.5	90	4	US-09-358-383C-31	Sequence 31, Appl
40	196.5	3.2	2353	4	US-08-984-709A-50	Sequence 50, Appl
41	184.5	3.0	1180	3	US-08-660-148-2	Sequence 2, Appl1
42	182	3.0	2273	4	US-09-426-998-5	Sequence 5, Appl1
43	181.5	3.0	1180	1	US-08-072-574-8	Sequence 8, Appl1
44	179.5	3.0	1180	1	US-08-466-270-8	Sequence 8, Appl1
45	179.5	3.0	1180	3	US-08-367-264-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1									
US-08-956-242-13									
: Sequence 13, Application US/08956242C									
: Patent No. 5986081									
: GENERAL INFORMATION:									
: APPLICANT: Gametxzy, Barry S.									
: TITLE OF INVENTION: Polynucleotides Encoding Herg-3									
: FILE REFERENCE: 960296.94550									
: CURRENT APPLICATION NUMBER: US/08/956,242C									
: NUMBER OF SEQ ID NOS: 13									
: SOFTWARE: PatentIn Ver. 2.0									
: SEQ ID NO 13									
: LENGTH: 1159									
: TYPE: PRT									
: ORGANISM: Homo sapien									
US-08-956-242-13									
Query Match									
Best Local Similarity 100.0%; Score 6079; DB 2; Length 1159;									
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MPVRGHVAPQNTFLDTIIRKFEQSRKFIITANARVENCAYICNDGFCELGYSRAEVM	60						
DB	1	MPVRGHVAPQNTFLDTIIRKFEQSRKFIITANARVENCAYICNDGFCELGYSRAEVM	60						
QY	61	ORPCTCFIAGPRTORRAAOIAQALGAERKVEIAFYRKDSCFICLVNVPVKNDG	120						
DB	61	ORPCTCFIAGPRTORRAAOIAQALGAERKVEIAFYRKDSCFICLVNVPVKNDG	120						
QY	121	AVIFIIILFEVYMEKDWGSPAHDTNHRGPTSWLABGRATPFLKLPALATLARESSV	180						
DB	121	AVIFIIILFEVYMEKDWGSPAHDTNHRGPTSWLABGRATPFLKLPALATLARESSV	180						
QY	181	RSGAGAGAGAVVVDLTPAAPSSSESLALDEVYAMDNHVAGLGAERKVEIAFYRKDSCFICLVNVPVKNDG	240						
DB	181	RSGAGAGAGAVVVDLTPAAPSSSESLALDEVYAMDNHVAGLGAERKVEIAFYRKDSCFICLVNVPVKNDG	240						
QY	241	PRSAAGLPSPRAHSLNDPDAAGSSCSLARTSRSCASVRRASSADDEARACVLP	300						
DB	241	PRSAAGLPSPRAHSLNDPDAAGSSCSLARTSRSCASVRRASSADDEARACVLP	300						
QY	301	RHAATGAMHPRLRSGLNSTSDSDLVYRTITKIPQITLNFVDLKGDFLASPTSDREITIA	360						
DB	301	RHAATGAMHPRLRSGLNSTSDSDLVYRTITKIPQITLNFVDLKGDFLASPTSDREITIA	360						
QY	361	PKIERHNVTAKTQVLSGADVLPEKILQAPRIHMTLHYSPPFAVMDLILLVIY	420						
DB	361	PKIERHNVTAKTQVLSGADVLPEKILQAPRIHMTLHYSPPFAVMDLILLVIY	420						


```

QY 421 TAVFTPYSAFLKTEEGPPATECGYACQPLAVVDLIVDIMEIVDILINFRTYVNAME 480
DB 421 TAVFTPYSAFLKTEEGPPATECGYACQPLAVVDLIVDIMEIVDILINFRTYVNAME 480
QY 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLFGSGSEELIGLLTARLLRVRARKID 540
DB 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLFGSGSEELIGLLTARLLRVRARKID 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIYVAINMEOPHMDSRIGMLHNLGDOIGKPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIYVAINMEOPHMDSRIGMLHNLGDOIGKPYNS 600
QY 601 GIGGSIIDKIVYVYALYFTFSSLTSGVFGVSPNTSEKIFSCVMLIGSLMTASIFGVNS 660
DB 601 GIGGSIIDKIVYVYALYFTFSSLTSGVFGVSPNTSEKIFSCVMLIGSLMTASIFGVNS 660
QY 661 AIORLYSGTARYHTOMLRVREFIRFHQIPNPLRORLEEFQHAWSYNGIDMNVKGF 720
DB 661 AIORLYSGTARYHTOMLRVREFIRFHQIPNPLRORLEEFQHAWSYNGIDMNVKGF 720
QY 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFTTTHAPGDTLVHAGDILLTAY 780
DB 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFTTTHAPGDTLVHAGDILLTAY 780
QY 781 FISRGSIELKGDVVAALGKNDIFGEPLNLYARPGKSGDVRLTYCDLKHIRDDLE 840
DB 781 FISRGSIELKGDVVAALGKNDIFGEPLNLYARPGKSGDVRLTYCDLKHIRDDLE 840
QY 841 VUDMYPEFSDHFWSSLEITFNLRODTNMTPGSPGSTELGSGFSROKRLKSPFRRTDKTE 900
DB 841 VUDMYPEFSDHFWSSLEITFNLRODTNMTPGSPGSTELGSGFSROKRLKSPFRRTDKTE 900
QY 901 QPGEVSALGPRAGAPSSRGRPGGPGWGESPPSGSPSSSEDEBPGRSSSPLRLVPPSS 960
DB 901 QPGEVSALGPRAGAPSSRGRPGGPGWGESPPSGSPSSSEDEBPGRSSSPLRLVPPSS 960
QY 961 PRPGEPPGGEPLMDECKSSDTCNPLSGAFSGVSNITFSFGDSRGRYOELPRCPAPTP 1020
DB 961 PRPGEPPGGEPLMDECKSSDTCNPLSGAFSGVSNITFSFGDSRGRYOELPRCPAPTP 1020
QY 1021 SLNLTPLSSPGRPRGVDRESRLDALOROLNLETPLSDMATAVLOLQROMTLVPAYSA 1080
DB 1021 SLNLTPLSSPGRPRGVDRESRLDALOROLNLETPLSDMATAVLOLQROMTLVPAYSA 1080
QY 1081 VTTPEPGPTSTPLPVSPPLTLILDSLSQVSOFMACELEPGABELPOEGFTTRLSPG 1140
DB 1081 VTTPEPGPTSTPLPVSPPLTLILDSLSQVSOFMACELEPGABELPOEGFTTRLSPG 1140
QY 1141 QLGALTSQPLHRHSGDPSG 1159
DB 1141 QLGALTSQPLHRHSGDPSG 1159

```

```

US-09-351-215-13
Query Match      100.0%; Score 6079; DB 3; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRRGHVAPONTFLDTIIRKTEGOSRKFTIANARVENCAYICNDGCELCGYSRAEYM 60
DB 1 MVRRGHVAPONTFLDTIIRKTEGOSRKFTIANARVENCAYICNDGCELCGYSRAEYM 60
QY 61 ORPCTDFLHGPRTORRAAQAQIAQLLGAEEKVEIAFRKQSGCELCVDDVVPKNDG 120
DB 61 ORPCTDFLHGPRTORRAAQAQIAQLLGAEEKVEIAFRKQSGCELCVDDVVPKNDG 120
QY 121 AVIMEILNEVYMEKDMGSPADPNHGRPPSWLAPGAKTFRLKPLALLATARESSV 180
DB 121 AVIMEILNEVYMEKDMGSPADPNHGRPPSWLAPGAKTFRLKPLALLATARESSV 180
QY 181 RSGGAGAGAPGAVVVDLTPAAPSSSELAIDEVTAMDNDHVAIGLPAERBALVPGSP 240
DB 181 RSGGAGAGAPGAVVVDLTPAAPSSSELAIDEVTAMDNDHVAIGLPAERBALVPGSP 240
QY 241 PSAPGOLPSPRAHSLNDPASGSSCLARTSRESCASVRRASSADDIEMARAGVLPPEP 300
DB 241 PSAPGOLPSPRAHSLNDPASGSSCLARTSRESCASVRRASSADDIEMARAGVLPPEP 300
QY 301 RHASTGAHPLRSGLNSTSDSLVRYRTISKIPOITLNFVDLKGDPPLASPTSDREIA 360
DB 301 RHASTGAHPLRSGLNSTSDSLVRYRTISKIPOITLNFVDLKGDPPLASPTSDREIA 360
QY 361 PKIKERTHNVTEKVOVSLGADVLEPKYKQAPRIHRTILHYSPEKAVMDLILLYIY 420
DB 361 PKIKERTHNVTEKVOVSLGADVLEPKYKQAPRIHRTILHYSPEKAVMDLILLYIY 420
QY 421 TAVFTPYSAFLKTEEGPPATECGYACQPLAVVDLIVDIMEIVDILINFRTYVNAME 480
DB 421 TAVFTPYSAFLKTEEGPPATECGYACQPLAVVDLIVDIMEIVDILINFRTYVNAME 480
QY 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLFGSGSEELIGLLTARLLRVRARKID 540
DB 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLFGSGSEELIGLLTARLLRVRARKID 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIYVAINMEOPHMDSRIGMLHNLGDOIGKPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIYVAINMEOPHMDSRIGMLHNLGDOIGKPYNS 600
QY 601 GIGGSIIDKIVYVYALYFTFSSLTSGVFGVSPNTSEKIFSCVMLIGSLMTASIFGVNS 660
DB 601 GIGGSIIDKIVYVYALYFTFSSLTSGVFGVSPNTSEKIFSCVMLIGSLMTASIFGVNS 660
QY 661 AIORLYSGTARYHTOMLRVREFIRFHQIPNPLRORLEEFQHAWSYNGIDMNVKGF 720
DB 661 AIORLYSGTARYHTOMLRVREFIRFHQIPNPLRORLEEFQHAWSYNGIDMNVKGF 720
QY 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFTTTHAPGDTLVHAGDILLTAY 780
DB 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFTTTHAPGDTLVHAGDILLTAY 780
QY 781 FISRGSIELKGDVVAALGKNDIFGEPLNLYARPGKSGDVRLTYCDLKHIRDDLE 840
DB 781 FISRGSIELKGDVVAALGKNDIFGEPLNLYARPGKSGDVRLTYCDLKHIRDDLE 840
QY 841 VUDMYPEFSDHFWSSLEITFNLRODTNMTPGSPGSTELGSGFSROKRLKSPFRRTDKTE 900
DB 841 VUDMYPEFSDHFWSSLEITFNLRODTNMTPGSPGSTELGSGFSROKRLKSPFRRTDKTE 900
QY 901 QPGEVSALGPRAGAPSSRGRPGGPGWGESPPSGSPSSSEDEBPGRSSSPLRLVPPSS 960
DB 901 QPGEVSALGPRAGAPSSRGRPGGPGWGESPPSGSPSSSEDEBPGRSSSPLRLVPPSS 960
QY 961 PRPGEPPGGEPLMDECKSSDTCNPLSGAFSGVSNITFSFGDSRGRYOELPRCPAPTP 1020
DB 961 PRPGEPPGGEPLMDECKSSDTCNPLSGAFSGVSNITFSFGDSRGRYOELPRCPAPTP 1020

```

```

RESULT 2
US-09-351-215-13
; Sequence 13, Application US/09351215
; Patent No. 6087488
; GENERAL INFORMATION:
; APPLICANT: Ganelzky, Barry S.
; TITLE OF INVENTION: Polynucleotides Encoding Herg-3
; FILE REFERENCE: 960296.94550
; CURRENT APPLICATION NUMBER: US/09/351,215
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: 08/956,242
; EARLIER FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapien

```

QY	1021	SLNIPILSSPPRRRGVESHDLAQRLNLEFLRLSDAATVQILQROMTLVPAYSA	1080
DB	1021		
QY	1081	VTTPGPGTSPILPVPSPILTLTLDLSISQSPMACEELPPGAPDELPOEGPTRLSLPG	1140
DB	1081		
QY	1141	QLGALTSQPLRHGSDPGS	1159
DB	1141		
		1141 QLGALTSQPLRHGSDPGS	1159
RESULT 3			
	US-09-226-012-2		
	; Sequence 2, Application us/09226012		
	; Patent No. 6207383		
	; GENERAL INFORMATION:		
	; APPLICANT: Keating, Mark T.		
	; APPLICANT: Spilaski, Igor		
	; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT		
	; TITLE OF INVENTION: SYNDROME GENE		
	; FILE REFERENCE: 2323-136		
	; CURRENT APPLICATION NUMBER: US/09/726,012		
	; EARLIER APPLICATION NUMBER: 09/122,847		
	; EARLIER FILING DATE: 1998-07-27		
	; NUMBER OF SEQ ID NOS: 116		
	; SOFTWARE: PatentIn Ver. 2.0		
	; SEQ ID NO 2		
	; LENGTH: 1159		
	; TYPE: PRT		
	; ORGANISM: Homo sapiens		
	US-09-226-012-2		
	Query Match	100.0%; Score 6079; DB 3; Length 1159;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 1159; Conservative	0; Mismatches	0; Indels
			0; Gaps
QY	1	MPVRGHVAPONTFLDTIIRKEGOSRKFTIINARVENCAYIYNDGFCELGYSRAEVM	60
DB	1	MPVRGHVAPONTFLDTIIRKEGOSRKFTIINARVENCAYIYNDGFCELGYSRAEVM	60
QY	61	QRPTCDPLHPRPQRRRAAQIAQALLGAERKVEIAFYRKDGSCFLCYDVYVYKNEG	120
DB	61	QRPTCDPLHPRPQRRRAAQIAQALLGAERKVEIAFYRKDGSCFLCYDVYVYKNEG	120
QY	121	AVIMEILNFEVYVMEKDMVGSPPAHNTNHRGPTSMAPRAKTFRLKPLPALLALARESSV	180
DB	121	AVIMEILNFEVYVMEKDMVGSPPAHNTNHRGPTSMAPRAKTFRLKPLPALLALARESSV	180
QY	181	RSGGAGAGADGAVVVDYLTLPAAFSSESLADEVTAMDNHVAGLPAEERRALVGPSP	240
DB	181	RSGGAGAGAGAVVVDYLTLPAAFSSESLADEVTAMDNHVAGLPAEERRALVGPSP	240
QY	241	PRSAAGOLPSRAHSLINDASGSSCSLARTSRESCASVRRASSADDTIEMARAGVLP	300
DB	241	PRSAAGOLPSRAHSLINDASGSSCSLARTSRESCASVRRASSADDTIEMARAGVLP	300
QY	301	RHASTGAMHPRSGSLNSTDSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIIA	360
DB	301	RHASTGAMHPRSGSLNSTDSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIIA	360
QY	361	PKIKERNHNTYEKTYOVLISGADVLPEYKIDAPRIHMTILHYSBFKAVMDMLILLVIY	420
DB	361	PKIKERNHNTYEKTYOVLISGADVLPEYKIDAPRIHMTILHYSBFKAVMDMLILLVIY	420
QY	421	TAVFPYSAAFILKTEEGRPATGCGYACOPLAVVDLVDIMPLIVDILINRTTYVANE	480
DB	421	TAVFPYSAAFILKTEEGRPATGCGYACOPLAVVDLVDIMPLIVDILINRTTYVANE	480
QY	481	EVSHSPRIAVHYEKGWELDMVAIPEDDLIFGSGSEELIGLKTARLLRVVARCLD	540

Db	Sequence	Score	DB 3	Length
Db	461 EVSHSPGIAAHYFEGWFLIDWVAIIPDDLLIFSGSBEELGLLKTAAFLVYARKKD	100.0%	DB 3	1159
QY	541 RYSEGAVALLELMCTFALLIHHMLACIYVAIGNNEQPMADSRIGMLHGLDOIKPYNS	100.0%		600
Db	541 RYSEGAVALLELMCTFALLIHHMLACIYVAIGNNEQPMADSRIGMLHGLDOIKPYNS	100.0%		600
QY	601 GLGSPKIDKYVYALYFTFSSLTYSVGFVNSPNTNSEKIFSIQVWLIGSLMYASIFGNVS	6079		660
Db	601 GLGSPKIDKYVYALYFTFSSLTYSVGFVNSPNTNSEKIFSIQVWLIGSLMYASIFGNVS	6079		660
QY	661 AIIQRLVSGTARYTQMLRVREFIRFHOIPMPLRORLEETVQHAMSTINGIDMAVYLGKF	720		720
Db	661 AIIQRLVSGTARYTQMLRVREFIRFHOIPMPLRORLEETVQHAMSTINGIDMAVYLGKF	720		720
QY	721 PECLOADIICLHUNSLNLOHCKPFGATGCRALAMKKTTHAPGDTLVHAGDLLTALY	780		780
Db	721 PECLOADIICLHUNSLNLOHCKPFGATGCRALAMKKTTHAPGDTLVHAGDLLTALY	780		780
QY	781 FISRGSIEILRGDVVAALIGKNDJFGEPLNLVYARPGKSGDVRLATYCDLHKIRDDLE	840		840
Db	781 FISRGSIEILRGDVVAALIGKNDJFGEPLNLVYARPGKSGDVRLATYCDLHKIRDDLE	840		840
QY	841 VLDWYPEESDHFWSLETFNLRDTNMIPGSPGSTELEGGSSRQRRKLSRRRTDKDTE	900		900
Db	841 VLDWYPEESDHFWSLETFNLRDTNMIPGSPGSTELEGGSSRQRRKLSRRRTDKDTE	900		900
QY	901 QPGEYSALCPGAGAGPSSRCRPGGPMGESPSSGSPSESEDEGPGRSSPLRLVPSS	960		960
Db	901 QPGEYSALCPGAGAGPSSRCRPGGPMGESPSSGSPSESEDEGPGRSSPLRLVPSS	960		960
QY	961 PRPEPEPPGGEPLMEDCEKSSDTCNPLSGAFSGVSNIFSPFGDSKRGROYELPRCPAPT	1020		1020
Db	961 PRPEPEPPGGEPLMEDCEKSSDTCNPLSGAFSGVSNIFSPFGDSKRGROYELPRCPAPT	1020		1020
QY	1021 SLNLTPLSSPGARRPGDVESRLDALQRLNLETRLASDMATVQLQLOROMTLYVPAYSA	1080		1080
Db	1021 SLNLTPLSSPGARRPGDVESRLDALQRLNLETRLASDMATVQLQLOROMTLYVPAYSA	1080		1080
QY	1081 VTTPEPGGTSTSPILPVSPPLTLTUDLSQVSOFAACELPPGAPELPQEGPTRRLSLPG	1140		1140
Db	1081 VTTPEPGGTSTSPILPVSPPLTLTUDLSQVSOFAACELPPGAPELPQEGPTRRLSLPG	1140		1140
QY	1141 QLGALTSQLHRHGSPPGS 1159			
Db	1141 QLGALTSQLHRHGSPPGS 1159			
RESULT 4				
US-09-226-012-4				
Sequence 4, Application US/09226012				
Patient No. 6207383				
GENERAL INFORMATION:				
APPLICANT: Keating, Mark T.				
APPLICANT: Spiwaski, Igor				
TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT				
FILE REFERENCE: 2323-136				
CURRENT APPLICATION NUMBER: US/09/226,012				
CURRENT FILING DATE: 1999-01-06				
EARLIER APPLICATION NUMBER: 09/122,847				
EARLIER FILING DATE: 1998-07-27				
NUMBER OF SEQ ID NOS: 116				
SOFTWARE: PatentIn Ver. 2.0				
SEQ ID NO: 4				
LENGTH: 1159				
TYPE: PRT				
ORGANISM: Homo sapiens				
US-09-226-012-4				
Query Match	100.0%	Score 6079	DB 3	Length 1159
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1159	Conservative	0	Mismatches	0
		0	Indels	0
			Gaps	0

QY 1 MPVRGHAAPONTFTDIIIRKEGOSRKFIINARVENCAYIYCDGFCCLGYSRAEYM 60
DB 1 MPVRGHAAPONTFTDIIIRKEGOSRKFIINARVENCAYIYCDGFCCLGYSRAEYM 60
QY 61 ORPCTCDLHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYDVVPYKNEBG 120
DB 61 ORPCTCDLHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYDVVPYKNEBG 120
QY 121 AVIMEFLNEEYVMEKDMGSPAHDTNHRGPTSMIAPGRAKTFRLKPLALLALTARESSV 180
DB 121 AVIMEFLNEEYVMEKDMGSPAHDTNHRGPTSMIAPGRAKTFRLKPLALLALTARESSV 180
QY 181 RSGGAGAGAPGAVVVDLTPAASSESLADEVTADNHNHAGIGPAERERALTVPKSP 240
DB 181 RSGGAGAGAPGAVVVDLTPAASSESLADEVTADNHNHAGIGPAERERALTVPKSP 240
QY 241 PRSAGOLPSPRAHSLNDPASGSSCSLARTRESRCASVRRASADDIEMAGVLPKPP 300
DB 241 PRSAGOLPSPRAHSLNDPASGSSCSLARTRESRCASVRRASADDIEMAGVLPKPP 300
QY 301 RHAATGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
DB 301 RHAATGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
QY 361 PKIKERTHNVEKTVQVLSIGADVLPEYKQAPRIHRMTILHYSFPAKAWMDLILLVY 420
DB 361 PKIKERTHNVEKTVQVLSIGADVLPEYKQAPRIHRMTILHYSFPAKAWMDLILLVY 420
QY 421 TAVFTPYSAFLKTEEBGPATTECGYACQPLAVVDLIVDIMEIVDILINFRTTYVNAME 480
DB 421 TAVFTPYSAFLKTEEBGPATTECGYACQPLAVVDLIVDIMEIVDILINFRTTYVNAME 480
QY 481 EVVSHRGRIAVHYFEGWFLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARAKLD 540
DB 481 EVVSHRGRIAVHYFEGWFLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARAKLD 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGWLHNIGDOIGRPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGWLHNIGDOIGRPYNS 600
QY 601 GUGGSIKDKYVTAIFYFTSSITSVGFVNSPNTSEKIFESICVWLIGSLAMSTIFGNVS 660
DB 601 GUGGSIKDKYVTAIFYFTSSITSVGFVNSPNTSEKIFESICVWLIGSLAMSTIFGNVS 660
QY 661 AIIOFLYSGTARVHTQMLRVREFIRFHQIPNPLRQLEBYFOHAMSNTGIDMNAVLKGF 720
DB 661 AIIOFLYSGTARVHTQMLRVREFIRFHQIPNPLRQLEBYFOHAMSNTGIDMNAVLKGF 720
QY 721 PECLADICLHNLNBSLQHKRFRGATKGLRALMKRTTTHAPRGDTLVHAGDILLTALY 780
DB 721 PECLADICLHNLNBSLQHKRFRGATKGLRALMKRTTTHAPRGDTLVHAGDILLTALY 780
QY 781 FTSRGSITELRGDVVVALIGKNDIFGEPLNTYARPGKSGDVRALTICOLHITHDDLE 840
DB 781 FTSRGSITELRGDVVVALIGKNDIFGEPLNTYARPGKSGDVRALTICOLHITHDDLE 840
QY 841 VLDMPPEFSDEHWSLEITFTNLRTNMTIPGSPSTELGEGFSRQKRKLSPFRRTDKDE 900
DB 841 VLDMPPEFSDEHWSLEITFTNLRTNMTIPGSPSTELGEGFSRQKRKLSPFRRTDKDE 900
QY 901 QPGEVSAIGPGRAGAPSSRGPRGPGWGESPPSSGPPSSDEDEGPRSSPLRLVPSS 960
DB 901 QPGEVSAIGPGRAGAPSSRGPRGPGWGESPPSSGPPSSDEDEGPRSSPLRLVPSS 960
QY 961 PRPPEEPGEPGLMDECKSSDTCNPLSGAFSGVSNITSPWGDSSGRQOELPRCPAPRP 1020
DB 961 PRPPEEPGEPGLMDECKSSDTCNPLSGAFSGVSNITSPWGDSSGRQOELPRCPAPRP 1020
QY 1021 SILNTPLSPPGRPRGADVESRLDALQROLNRLFTLSDMAIVLQLOKQMLVVERPASA 1080
DB 1021 SILNTPLSPPGRPRGADVESRLDALQROLNRLFTLSDMAIVLQLOKQMLVVERPASA 1080

QY 1081 VTPBPGPTSTSPLLPVSPPLPTLTIDSLQVSQFMACEELPRGABELPOEGPTRLSLPG 1140
DB 1081 VTPBPGPTSTSPLLPVSPPLPTLTIDSLQVSQFMACEELPRGABELPOEGPTRLSLPG 1140
QY 1141 OUGALTQPLHRHSGSDPS 1159
DB 1141 OUGALTQPLHRHSGSDPS 1159
RESULT 5
US-09-358-383C-10
; Sequence 10, Application US/09358383C
; Patent No. 6518398
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy A.J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-055CP
; CURRENT APPLICATION NUMBER: US/09/358, 383C
; PRIOR FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-358-383C-10
Query Match 100.0%; Score 6079; DB 4; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPVRGHAAPONTFTDIIIRKEGOSRKFIINARVENCAYIYCDGFCCLGYSRAEYM 60
DB 1 MPVRGHAAPONTFTDIIIRKEGOSRKFIINARVENCAYIYCDGFCCLGYSRAEYM 60
QY 61 ORPCTCDLHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYDVVPYKNEBG 120
DB 61 ORPCTCDLHGPRTORRAAOIAQALLGAERKVEIAFYRKDGSCFLCYDVVPYKNEBG 120
QY 121 AVIMEFLNEEYVMEKDMGSPAHDTNHRGPTSMIAPGRAKTFRLKPLALLALTARESSV 180
DB 121 AVIMEFLNEEYVMEKDMGSPAHDTNHRGPTSMIAPGRAKTFRLKPLALLALTARESSV 180
QY 181 RSGGAGAGAPGAVVVDLTPAASSESLADEVTADNHNHAGIGPAERERALTVPKSP 240
DB 181 RSGGAGAGAPGAVVVDLTPAASSESLADEVTADNHNHAGIGPAERERALTVPKSP 240
QY 241 PRSAGOLPSPRAHSLNDPASGSSCSLARTRESRCASVRRASADDIEMAGVLPKPP 300
DB 241 PRSAGOLPSPRAHSLNDPASGSSCSLARTRESRCASVRRASADDIEMAGVLPKPP 300
QY 301 RHAATGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
DB 301 RHAATGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
QY 361 PKIKERTHNVEKTVQVLSIGADVLPEYKQAPRIHRMTILHYSFPAKAWMDLILLVY 420
DB 361 PKIKERTHNVEKTVQVLSIGADVLPEYKQAPRIHRMTILHYSFPAKAWMDLILLVY 420
QY 421 TAVFTPYSAFLKTEEBGPATTECGYACQPLAVVDLIVDIMEIVDILINFRTTYVNAME 480
DB 421 TAVFTPYSAFLKTEEBGPATTECGYACQPLAVVDLIVDIMEIVDILINFRTTYVNAME 480
QY 481 EVVSHRGRIAVHYFEGWFLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARAKLD 540
DB 481 EVVSHRGRIAVHYFEGWFLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARAKLD 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGWLHNIGDOIGRPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGWLHNIGDOIGRPYNS 600


```
FEATURE:
NAME/KEY: unsure
LOCATION: (637)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (646)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (653)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (656)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (660)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (661)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (662)
OTHER INFORMATION: Unidentified at time of filing
US-08-956-242-4
```

Query Match 40.1%; Score 2436.5; DB 2; Length 888;
Best Local Similarity 57.4%; Fred. No. 9.1e-197;

Matches 528; Conservative 71; Mismatches 166; Indels 155; Gaps 20;

```
QY 358 IIAKIKERTHNVEKYOVSLGADVPEYKLOAPRHRWTLIHSYSPKAWMDLILL 417
DB 1 IIAKIKERTHNVEKYOVSLGADVPEYKLOAPRHRWTLIHSYSPKAWMDLILL 60
QY 418 VIYAVPTPYSAAPFLKTEEGPATECGYACQPLAVVDLIVDIMFIVDLINERTYVN 477
DB 61 VIYAVPTPYSAAPFLKTEEGPATECGYACQPLAVVDLIVDIMFIVDLINERTYVN 119
QY 478 ANEEVSHRGLAVHYEFGWFLIDVVAIPFDLIFGSGSEE--LIGLLTARLLRVR 534
DB 120 QNEEVSPDPAIAHYEFGWFLIDVVAIPFDLIFGSGSDETTLLIGLLTARLLRVR 179
QY 535 VARKIDRYSEYGAALFLMCTFALIAHMLACIWAICNMEOBHMDSRIGLHNLGDOIG 594
DB 180 VARKIDRYSEYGAALFLMCTFALIAHMLACIWAICNMEOBHMDSRIGLHNLGDOIG 239
QY 595 KPYN-SSGLGSPSIRKDYVYALYFTSSLSVGFQVNSPNTNSEKIFSICVWLIGSLMYA 653
DB 240 KRYNDSOSSGSPSIRKDYVYALYFTSSLSVGFQVNSPNTNSEKIFSICVWLIGSLMYA 299
QY 654 SIFGNVSAIIORLYSGTARYHTOMRVREFIRPHOIPMLRORLEEYOHAMSYNGIDM 713
DB 300 SIFGNVSAIIORLYSGTARYHTOMRVREFIRPHOIPMLRORLEEYOHAMSYNGIDM 359
QY 714 NAVLKGPEECLOADICHLNLSLLOHCKPFRGATKGCRLALAMKFTNAPGDTLVHAG 773
DB 360 NAVLKGPEECLOADICHLNLSLLOHCKPFRGATKGCRLALAMKFTNAPGDTLVHAG 419
QY 774 DLTAALYFISGSIETLIGDVVAALIGKNDIFGEPLNLVYARPGNSNGDVALTYCDLHKI 833
DB 420 DLTAALYFISGSIETLIGDVVAALIGKNDIFGEPLNLVYARPGNSNGDVALTYCDLHKI 479
QY 834 HRDLEVLDMYPERSDHFMSSLETFEPLRDTNM---IPGSPGSELEGGSSROKRRKLS 890
DB 480 QREDDLEVLDMYPERSDHFMSSLETFEPLRDTNM---IPGSPGSELEGGSSROKRRKLS 539
QY 891 FRRRTDKD--TEOPGE-VSALGPRGAGPSSRGPGGPMGESP-----SSGSPSP 938
DB 540 FESGGEKENSTNDEPDSADTIRHVOSSXXPLKRXKXDPPLXSPPLAMKXLFSLVDSX 599
```

```
QY 939 ESSDEGFGRSSPLRLVFPSSPRPPGEPPEGLMEDCEKSSDTCNPLSGAFSGVSNIF 998
DB 600 XIGESAG-----LXFXXTCPPPGE-----CPYIXXPPLAXXS--- 630
QY 999 SFWGDSRGRYOXOELPR-----CPA-----PPPSLINTPL----- 1027
DB 631 --WLPYLXGK--NSLPRLXEPFLCPSPXNCXGFLXXVPTSGRMHIDKRSKCDITDMRWS 687
QY 1028 -----SPGRRPR-----GDVESRLDALORNLRLTSLADMAT 1062
DB 688 ERENHNPREDSSPSALORAMGISETSDDLTYGEBQRDLLOLQNLRLSQTMTDIOF 747
QY 1063 VLQLLQROMTLVPPASVATTPGPGPTSPPL-----PVSPLPT--LTLDSLSQVSQF 1114
DB 748 ILQLLQKOTVVPAPASVMTA---GSEYORPILQIMRTSOPASIKTDSPSPSSQCEPF 804
QY 1115 MACCELPPGAPLPOEG-----PVRRLSLP 1139
DB 805 LDLEKSKLSKSESLSGVHLNTASEDNLTSLKODSDLSLEHLRQRTYVVP1RHPSLP 864
QY 1140 -GOLGALTSOPLHRRGSDPG 1158
DB 865 DSSLSTVGIVGLHRRVSDPG 884
```

```
RESULT 7
US-09-351-215-4
Sequence 4, Application US/09351215
Patent No. 6087488
GENERAL INFORMATION:
APPLICANT: Ganetzky, Barry S.
APPLICANT: Titus, Steven A.
TITLE OF INVENTION: Polynucleotides Encoding Herg-3
FILE REFERENCE: 960296.94550
CURRENT APPLICATION NUMBER: US/09/351,215
EARLIER APPLICATION NUMBER: 08/956,242
EARLIER FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 4
LENGTH: 888
TYPE: PRT
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: unsure
LOCATION: (181)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (438)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (439)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (567)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (568)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (571)
OTHER INFORMATION: Unidentified at time of filing
```

FEATURE:
NAME/KEY: unsure
LOCATION: (575)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (580)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (589)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (593)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (599)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (600)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (608)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (610)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (611)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (621)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (624)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (625)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (628)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (629)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (634)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (637)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (646)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (653)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:

NAME/KEY: unsure
LOCATION: (636)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (660)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (661)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (662)
OTHER INFORMATION: Unidentified at time of filing
US-09-351-215-4

Query Match 40.1%; Score 2436.5; DB 3; Length 888;
Best Local Similarity 57.4%; Pred. No. 9.1e-197;
Matches 528; Conservative 71; Mismatches 166; Indels 155; Gaps 20;

Qy 358 IIAPIKERTHNWTEKTVOLSLGADVLPEXKLOAPRIHMTILHSPKAWMDLILL 417
Db 1 IIAPIKADRTNHWTEKTVOLSLGADVLPEXKLOAPRIHMTILHSPKAWMDLILL 60
Qy 418 VIYAVETPYSAFLKETEGBEPATECGYACOPLAIVDLIVDILINFRTYVN 477
Db 61 VIYAVETPYSAFLKETEGBEPATECGYACOPLAIVDLIVDILINFRTYVN 119
Qy 478 ANERVSHPCRIAVHYKGFLLDMVAALPDDLIFSGSSE--LGLKTAFLRLVR 534
Db 120 QNEEVSDPAKIAHYKGFLLDMVAALPDDLIFSGSSE--LGLKTAFLRLVR 179
Qy 535 VARKLDRTSEYGAVALFLMCTFALIAHMLACIYALGNMOPHMSRIGLHNGOIG 594
Db 180 VARKLDRTSEYGAVALFLMCTFALIAHMLACIYALGNMOPHMSRIGLHNGOIG 239
Qy 595 KPVN-SSGLGPPSIRKDYVVALYFTFSSLSVGVGNVSPMTNSEKIFSCVMLGSLMYA 653
Db 240 KRVNDSSSSSPSIRKDYVVALYFTFSSLSVGVGNVSPMTNSEKIFSCVMLGSLMYA 299
Qy 654 SIFGNVSAIIORLYSGARHTOMLRYREFRFOIENPLRQRLSEYFOHMASTNGIDM 713
Db 300 SIFGNVSAIIORLYSGARHTOMLRYREFRFOIENPLRQRLSEYFOHMASTNGIDM 359
Qy 714 NAVLKGPPECLADICHLNRSILHOCKPPFAGTKGCLRALAMFKTTTHAPGDTLVHAG 773
Db 360 NAVLKGPPECLADICHLNRSILHOCKPPFAGTKGCLRALAMFKTTTHAPGDTLVHAG 419
Qy 774 DLATLALYISNGSIELRGDVVAALIGKNDIFGEPLNLVYAPGKSGNDVVALTYCDLHKI 833
Db 420 DLATLALYISNGSIELRGDVVAALIGKNDIFGEPLNLVYAPGKSGNDVVALTYCDLHKI 479
Qy 834 HRDDLLEVLDMYPPFSHFSSLEITFNLRDTNM---IPSGSGTELEGESRORRKL 890
Db 480 HRDDLLEVLDMYPPFSHFSSLEITFNLRDTNM---IPSGSGTELEGESRORRKL 539
Qy 891 FRRRTDRD--TEOGE-VSALGPGRAGSPSSRGRPGPMGESP-----SSGSSP 938
Db 540 FRRRTDRD--TEOGE-VSALGPGRAGSPSSRGRPGPMGESP-----SSGSSP 599
Qy 939 ESSEDEGPRGSSPLRLVPPSSPRPGEPPGEPJLMDCEKSDTCNPLSGAFSGVSNIF 998
Db 600 ESSEDEGPRGSSPLRLVPPSSPRPGEPPGEPJLMDCEKSDTCNPLSGAFSGVSNIF 630
Qy 999 SFMGDSRGRYOELPR-----CPA-----PTPSLNLPL----- 1027
Db 631 --WLPXLGXK-NSLPLRXPLCPXKXGFLXXXVPTSGMHLDKRSHSCDKITDMRSW 687
Qy 1028 -----SSPGRBPR-----GDVESRLDALORLNRLLETRLSADMAT 1062
Db 688 ERENAPQPEDSSSALQRAAWGISFTESDLITGEVORLDLLOEQLNRLLESQMTTIO 747

Db 138 -----VKTP-----EDKEDKVK----- 151
QY 240 PPRSAPQLSPRAHSLNPDASGSSCSLARTSRSCASVARRASSADDIEMRAGVLP 299
152 -----GRSRACTHDSARRRRA----- 169
QY 300 PRAHSTAGMHLRSLNSTSDLVKRTYSKIPOJTLNPFVLDKGFPLASPSDREII 359
170 -----VLYHISGHLQRRKKNLKNNNVFDKP----- 197
QY 360 APRIKERTHNTEKVTQVLSGADVLPDEYKLOAPRIHRMTLHSPKRAVMDLILLVI 419
198 -----APPEYKVSDAKSKSFILHSTFPAQMDWILLATF 233
QY 420 YTAVFTEYSAFLLKETEERGPATECGYACOPLAVDLIDIMFIVDILNFRTTYVAN 479
234 YVAATVBYNCFIGNDLSLTRST-----TVSDIAVEILFDILNFRTTYYSKS 284
QY 480 EEVYSHFGRIAVHFKMFLIDMVAIIPDLL-IFGSGSEELIGLKTARLRLVYARK 538
285 GGVTFEARSICIHVYTWTFIIDLIALPFDLLAFNVTYVSLVHLKTVRLRLRLQK 344
QY 539 LDRYSEGAVALFLMCTFALIAHMACIWAIGNMOPH--MDSRIQWLNLDQIGK 595
345 LDRISQSTIVLILMSFALLAHMACIWTYIGKMEREDNSLTKWEYGMJHEGKRLS 404
QY 596 P-YNNSGLGSPSICKDYVYALYFTFSSILTVGFGNVPSPNTSEKIFISICVALISLWYAS 654
405 PYYGNNTLGGPSISAVIAIALYFTLSLTVSGFGNVSANTDAEKIFISICTLIALMAL 464
QY 655 IFGNVSAIIQRLYSGTARVHTOMLRVREFIFHOIIPMLRORLEEYFOHANSYNGIDMN 714
465 VFGVNTAIQMYRMSLYHTRTKDLDFIRVHHLPOOLKORMLEYEQTTSVYNGIDSN 524
QY 715 AVLKFPEDLOADICLHNSLSLOHCKPFGATGKCLRALAMKFTTHAPRGDVLVHAG 774
525 ELKDPDELRSDDITMHLNKEIID-LSLFECASRGCLSLSLHTIKTSCAFGEYLLROGD 583
QY 775 LITLALYISRGSIIILGDDVVALIGKNDIFGEPFLNLYARPGKSGVDRALTYDCLKIH 834
584 ALQAIYFVCGSMELVLDMSVLAITLKGDDILGANLSIKDOYIKTNADVKALTYDLOCII 643
QY 835 RDOLLEVLDMYPEPSDHFWSL--EITFNLND-----TNMIPGSPGSTELEGFSQOR 885
644 LKGFLEVGLDYPEYAHKFVEDIQHDLTYNLRGHESDVISRLSNKSNVSOSEPGNGNIN 703
QY 886 KRKLSFRRTDKDIOEG--EVSALGPRAGAGPSSSRGPRG-----PWGE 929
704 KRLPISYE--DEEEEBCEEEANVSLP-ICTRGSSSRNKKVSGSKAYLGLSLQOLASGT 760
QY 930 SPSSGP-----SSPESSEDEGPGRSSS-----PLRLVPPSPRPPEPPGEPPLMEDC 977
761 VPFHSPIRVSRNSPKTKOEIDPRNNHKKREKNLKLQSLTNMAGPDLSPRIYDGTEDG 820
QY 978 EKSSDTCNPLSGASGVSNIFSPWGDGRGOYOLPKCPATPPLNIPLSSPGRRRGD 1037
821 NSSE-----SQTFDF-GSBRIR-----SEPRISP---PLGD 848
QY 1038 VE-----SRDLAOLQNLRL-----TRLSDMATVLOLOROMTLPVPA 1077
Db 849 PEIAAVALFIKAETKQOINKLNSEVITTLQOEVSQLGKDMKVNIRLEENLVSPQPS 905

PRIOR APPLICATION NUMBER: USSN 09/119,855
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 1102
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-358-383C-36
Query Match 25.5%; Score 1551; DB 4; Length 1102;
Best Local Similarity 32.6%; Pred. No. 8.3e-122;
Matches 383; Conservative 169; Mismatches 283; Indels 340; Gaps 30;
QY 1 MPVRGHVAPQNTFLDTIIRKFEQSGKFIANARV-ENCAYIYNDGFCGLCYSAEV 59
1 MPVAKGLLPQNTFLDTIATRFQDTHSNFIANQVAKGPFIYVCSDFCLAGFAREV 60
Db 60 MQRPTCDLHGPRTPORRAAQIQAALLGAEEKVEYLAIFYKKDSCFLCYDVYVKNED 119
61 MOKSCSKFLFGEVETNQLMQLIEKSLSEKVEFGEIMFYKKNGAPFWCLDIVIPINKEK 120
QY 120 GAVIMELINEFVMEKMDVGSPPADTNHRGPTSMLAGRAKTRFLKLPALLATARESS 179
121 GDVYVLFASF-----KDTITIKV----- 138
Db 180 VRSGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAAGLPAERRALVGPSS 239
139 -----KISEDK-----KEDR----- 149
QY 240 PPRSAPQLSPRAHSLNPDASGSSCSLARTSRSCASVARRASSADDIEMRAGVLP 299
150 ---AKGR-----SRAGHPDSARRSR----- 168
QY 300 PRAHSTAGMHLRSLNSTSDLVKRYRTYSKIPOJTLNPFVLDKGFPLASPSDREII 359
169 -----AVLYHISGHLQRRKKNLKNNNVFDKP----- 197
QY 360 APRIKERTHNTEKVTQVLSGADVLPDEYKLOAPRIHRMTLHSPKRAVMDLILLVI 419
198 -----APPEYKVSDAKSKSFILHSTFPAQMDWILLATF 233
QY 420 YTAVFTEYSAFLLKETEERGPATECGYACOPLAVDLIDIMFIVDILNFRTTYVAN 479
234 YVAATVBYNCFIGNDLSLTRST-----TVSDIAVEILFDILNFRTTYYSKS 284
QY 480 EEVYSHFGRIAVHFKMFLIDMVAIIPDLL-IFGSGSEELIGLKTARLRLVYARK 538
285 GGVTFEARSICIHVYTWTFIIDLIALPFDLLAFNVTYVSLVHLKTVRLRLRLQK 344
QY 539 LDRYSEGAVALFLMCTFALIAHMACIWAIGNMOPH--MDSRIQWLNLDQIGK 595
345 LDRISQSTIVLILMSFALLAHMACIWTYIGKMEREDNSLTKWEYGMJHEGKRLS 404
QY 596 P-YNNSGLGSPSICKDYVYALYFTFSSILTVGFGNVPSPNTSEKIFISICVALISLWYAS 654
405 PYYGNNTLGGPSISAVIAIALYFTLSLTVSGFGNVSANTDAEKIFISICTLIALMAL 464
QY 655 IFGNVSAIIQRLYSGTARVHTOMLRVREFIFHOIIPMLRORLEEYFOHANSYNGIDMN 714
465 VFGVNTAIQMYRMSLYHTRTKDLDFIRVHHLPOOLKORMLEYEQTTSVYNGIDSN 524
QY 715 AVLKFPEDLOADICLHNSLSLOHCKPFGATGKCLRALAMKFTTHAPRGDVLVHAG 774
525 ELKDPDELRSDDITMHLNKEIID-LSLFECASRGCLSLSLHTIKTSCAFGEYLLROGD 583
QY 775 LITLALYISRGSIIILGDDVVALIGKNDIFGEPFLNLYARPGKSGVDRALTYDCLKIH 834
584 ALQAIYFVCGSMELVLDMSVLAITLKGDDILGANLSIKDOYIKTNADVKALTYDLOCII 643
QY 835 RDOLLEVLDMYPEPSDHFWSL--EITFNLND-----TNMIPGSPGSTELEG 880
644 LKGFLEVGLDYPEYAHKFVEDIQHDLTYNLRGHESDVISRLSNKSNVSOSEPGNGNIN 703

```

QY 881 !SRQKRLKSRRTDXDTEQPGVSLALGPRAAGASRCRRPGGPMGCESSSGSPSES 940
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 700 GS-!KKLPLSTVEDDEEEVEEEETTISIP-----!YTRK-----SSVSH 738
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 941 SEDEGPGRSSPLRL-----VFSSP-----RPGGEP-----968
    !::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 739 SKRTGSSKSYGLSLKQLTSGTVPFHSPRIYSSANSPKTKQADPNNHGTREKKNLKVQL 798
    !::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 969 -----GGEPLM-----EDCEKSSDPCNPLSAGFSVSNISFSGMDSRKRQVQLPRC 1015
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 799 CSLCTAGAPELSPRIVDQIEDGNSSEET-----OTFFGSE---QIRPEPRI 842
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 1016 PAPPTSLNIPILSSPGRRPGDVE-----SRDLQRLNRL-----TRLSA 1058
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 843 ---SPSL-----GESLGAFAFLIKAEFRKQQLNKLNSEVTTILQEVSOQIK 886
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 1059 DMATVQLQLQKMTLVPPAYSAVTTPCGDPPSTSP 1093
    ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 887 DMRISIMOLENLLISPOQPSQPC---SLHPSTQPC 917
    ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 13
US-09-343-494-1
; Sequence 1, Application US/09343494
; Patent No. 6413741
; GENERAL INFORMATION:
; APPLICANT: Jegla, Timothy J.
; APPLICANT: Wickenden, Alan
; APPLICANT: ICGen, Incorporated
; TITLE OF INVENTION: Human Elk, a Voltage-Gated Potassium Channel Subunit
; FILE REFERENCE: 0.18512-001320US
; CURRENT APPLICATION NUMBER: US/09/343,494
; EARLIER FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: US 60/091,469
; EARLIER FILING DATE: 1998-07-01
; EARLIER APPLICATION NUMBER: US 60/116,621
; EARLIER FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1083
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human Elk (hElk; Eag (eather a go-go)-like K+ gene)
; OTHER INFORMATION: potassium channel monomer protein
US-09-343-494-1

```

[illegible]

QY	300	FRHASTGAMHDLRSGLLNSTSDSLVRRRTISKIPQITLNNVVDLKGDFPLASPTSDKEII	359
Db	185	PK-----GKHLNKGVEGEPKN-----	201
QY	360	APKIKERTHNYTEKVQVLSGADVLEPYKLOAPRIHRWTLIHXSPKAVMDLILLVI	419
Db	202	-----LPEYKAAAIRKSPFILLHCGALRAVWDGILLATL	236
QY	420	YTAFTPYSAFLKTEEGPPATECGYACOPLAVVDLIVDIMEITVDILINFRTYYNAN	479
Db	237	YVAATVPYSVCV-----STAREPSA-----ARGPSPVCDAVEVLEITLIVLNFRTFVSKS	288
QY	480	EEVYSHPORIAVHYEKGMELIDMYAAIPEDL-----IFGSSSEELIGLTKTARILL	532
Db	289	GOVVPAPKSPICLHYVYTMFLVDVLAALDEPFIDLLHAKVNYVFGA-----HLTKVRLRL	342
QY	533	VRVAKRIDRYSEYGAVALFELMCFALIAHMLACIWAIGNMEQPHMS--RIGWMLNL	589
Db	343	LRLPRIDRYSQYSAVVTLLMANFALLAHNVACVWFYIGORELSESESELPELGMQEL	402
QY	590	GDQIGKPY-----NMSG-----	621
Db	403	ARRLPETPYVILGRPPAGCNSSGQSDNCSSSSEANGCTGELLGCSLISAYTISLYFALSS	462
QY	622	LTSVGFQVNSPTNSEKFEFICWVLIGLSIMASTFQVNASVALIOBYSGARVHOMLRVR	681
Db	463	LTSVGFQVNASNTDIEKTFESTCITLIGLMAHVAVFGVNTALTQOMVYARFLYBRTDOLR	522
QY	682	EFIRFHOIPLNRLORLEEFQHAHSYTGIDMANAVLKGFPICLOADICLHLNRSLOHCK	741
Db	523	DYIRHRIKPKLQKQMLEYFOATMAVNNGIDTTELOSLPELADIAHMLHKVLDL-LP	581
QY	742	PFKQATKCCALALAMKEFTTHAPRGDTLVHAGDILLTALYFISRSIELKGDVVAIILG	801
Db	582	LFEASRCCIALSLALMPACETGEXYLIHGDALQALYFVCSGSMELKGGVLAIIIGK	641
QY	802	NDIFGEPLNLAIRPGKSGNGVRAITYCDLHKIHRDDLEVDLMYPEESDHFWSSL--EIT	859
Db	642	GDILGCELPRRBOYYKANADYKGLTFYCYLOQLAGLHDSIALXPFAPARSRBOLRGELS	701
QY	860	FNLRDNTMI PGSPGSTELEGFSFQKRRKUSFRRTDKDYEQ-----	901
Db	702	YNL-----GAGGCGAEVDYTSLSGDNTLMSTLEEKETDGBQGPVSPAPADESSPLLS	755
QY	902	PGEVSA-----LGGGRAGAP--SSRRPGPMGSESPSSPESSEDEGPPRRSSPL	953
Db	756	PGCTSSSSAAKLISRRTPAPRPLUGRGRRPAGALKAEAQSPAP-----RALEGL	807
QY	954	RLVPPSSRPGEPPGEGEP-LMEDCEKSSDPCNLSGAFSGVSNISFPMWDSRROYOEL	1012
Db	808	RL-----PPMPNVPDLSPRYVDGIEDGCCGDQK-----FSF-----RVQSG	847
QY	1013	PRC-----PAPTPSLINIPLSSPGRRPGVDYER--LDALQOLNRLBETRLSADMATV	1063
Db	848	PECSSSPPGPEGSGILLTV-----PHGPPEARWTDPLDKLROAVTELSEOV-LQHREG	898
QY	1064	LQILQORQTLV-----	1082
Db	899	LQSLKQAOVLVLAHREBPCPRASGEGPCPASTSGLLQPLCVDYGASSYCLQPPAGSVLS	958
QY	1083	-----TVGPPTST-----SPLLVPSPLP-----TLTDLISQVSOQFMACEPLPGAP	1125
Db	959	GTWPHAPAGPPPLMAPMPWGPSPAGSSWPMPATAFMWSTSDSEPPASGDCLSESTPASP	1018
QY	1126	ELPOEG 1131	
Db	1019	PPSEEG 1024	

RESULT 14
US-09-600-776-2
; Sequence 2, Application US/09600776
; Patent No. 6326168


```

QY 180 VRSGAGAGAPCAVVVDVLTTPAAPSSSESLALDEVATMDNHVAGLGAPEERRALVPGS 239
   146 -KETSGS-----RRR-----154
QY 240 PPSAPGOLPPRAHSLNPDASGSSCLARTSRRESCASVRRASADDIEFAMAGVLP 299
   155 -----YGRARSKGFNANRRRSRA--VLYHLSGHLQK 184
QY 300 PRAHSGAMHPLASGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIT 359
   185 PK-----GKHKLKGVFGEKPN-----201
QY 360 APRKERTHNTEKYQVLSLADVLPEYKLOAPRIHRMTILHSPPKAVMDWLLLVY 419
   202 -----LPEYVAAIRKSPFILLHCGALRATWDGFIILLATL 236
QY 420 YTAFTPYSAFLLKTEEGPPATECGYACQPLAVVDLIVDIMEFVIDILINFRTYVNAV 479
   237 YVAVTPYVCY--STARPSA-----ARGPPVCDLAVEVLFIDIVLNFRTFVSKS 288
QY 480 EEVYSHRGRIAVYFKGFIDMVAIPDIL-----IFSGSEELIGLTKTARLLRL 532
   289 GQVVFAPKSTICLHYVTTFLLDVIYALPFDLHAFKVNYFEA-----HLKTVRLRL 342
QY 533 VRVARKLDYRSEYGAAVFLMCTFALIAHMLACIWAYAIGNMOPHMDS---RIGWHLNL 589
   343 LRLPLRDRYSQYSAVVLTLMAVFLAHVACWFIYIGOREISESESELPEIGWLOEL 402
QY 590 GDOIGKPY-----NSSG-----LGGPSIKDKYVYALYFTSS 621
   403 ARRLTEPYLVGRRPAGNNSGQSDNCSSSEANGTGLLGGPSLRSAYITSLYFALSS 462
QY 622 LTVSGGNVSPNNTSEKIFESICVMLIGSLMYASIFGVNSAIIQRIYSGTARHTOMLRVR 681
   463 LTVSGGNVNSANDTEKIFESICVMLIGALMAHVAFFGNVTALIQRYARFLYHSRTRDLR 522
QY 682 EFTIRFOIPNPLRQRLSEYFOHMASYTINGIDMNAVLTGKPEECLOADICILNRSLLQHK 741
   523 DYIRIRIRPKRLQRMLEYQATWAVNNGIDITTELOSLEPDELRAIDIAHLHKEVLO-LP 581
QY 742 PFRGATKGLRALAMFKTTTHAPPGDTLVHAGDULTALYFISRGSIEILRGDVVAAILGK 801
   582 LFEASRGLRALSLRALPAPCTPGEYLIHQDALQALYFVCSGMEVLKGVYLAITLGK 641
QY 802 NDIFFGEPLNIYARPGKSNQVRAITCYCDLKH IHRDILLEVLDMYPEPSDHFWSL--EIT 859
   642 GDILGCELPREQYVVKANADVKGITYCVLQCLQLAGLHDSLALYPEFAPRFSRGLRGELS 701
QY 860 FNLBDTMMITGSPGSTLEGGFSRQRRKLSFRRTDKDTEQ-----901
   702 YNL-----GAGGGSALVDYTSLSGDNLTLSLEEKETDGEQPTVSPAPADESSPLLS 755
QY 902 PGEVSA-----LGPGRAGAP--SSRGRPGPGWGESPPSSPSESSEDEGPGRSSPL 953
   756 PGCSSSSSAKALSPRTAPRPRGGRGRGAGALKAEGAPSA-P-----RALEGL 807
QY 954 RLYPFSSPPRPGEPGEP--LMEDEKSSPTCNPLSGAFSGVSNIFSFMGDSRGROYQEL 1012
   808 RL-----PPMPWNVPPDLSPRVVDGIEDGCGSDQPK-----FSF-----RMQOSG 847
QY 1013 PRC-----PAPTPSLNIPILSSPGRRRRGVRESR---LDALQRLRLLETRLSDMATV 1063
   848 PECSSSPSPGESGLTV-----PHGPEARNITDTLKLQAVMELSEQV-LQMRG 898
QY 1064 LQLLQROMTLV-----PPAYSAYT 1082
   899 LQSLRQAVOLVLANHREGPCPRASGEGPCPASTSGLLOPLCVDTGASSYCLQPPRAGSVLS 958
QY 1083 -----TPGPGPTST-----SPLLVPSPPLTLTIDSLSQVQFMACEELPPGAPELDPEG 1131
   959 GTPWPHRRPBPRLAPWPMGCPRASQSSPMPRATAFMTS-----TSDSEPPASGDLCSGP 1012

```

```

QY 1132 PTRRLSLPGQLGALTSOP 1149
Db 1013 STPASPPSEEGARTGPP 1030

```

Search completed: September 23, 2003, 17:41:04
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 17:37:05 ; Search time 455 Seconds
(without alignments)
379.269 Million cell updates/sec

Title: US-10-000-151B-3
Perfect score: 6079
Sequence: 1 MPVRGHVAPONTFLDTIIR.....COLGALTSQPLRHGSDPGS 1159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 556269 seqs, 148893369 residues

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Published_Applications_AA.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCITUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6079	100.0	1159	9	US-09-735-995-2	Sequence 2, Appl1
2	6079	100.0	1159	9	US-09-735-995-10	Sequence 4, Appl1
3	6079	100.0	1159	9	US-09-119-855-14	Sequence 10, Appl1
4	6079	100.0	1159	15	US-10-000-151B-3	Sequence 5, Appl1
5	6079	100.0	1159	15	US-10-193-692-5	Sequence 10, Appl1
6	6079	100.0	1159	15	US-10-185-867-10	Sequence 49, Appl1
7	6079	100.0	1159	15	US-10-174-613-49	Sequence 2, Appl1
8	3493.5	57.5	1177	15	US-10-193-692-4	Sequence 4, Appl1
9	3312.5	54.5	1195	15	US-10-192-440-10	Sequence 10, Appl1
10	3255	53.5	1186	15	US-10-255-532-2	Sequence 2, Appl1
11	3241	53.3	1186	15	US-10-193-692-2	Sequence 8, Appl1
12	3043.5	50.1	994	15	US-10-325-430-21	Sequence 21, Appl1
13	3001.5	49.4	958	12	US-10-325-430-21	Sequence 2, Appl1
14	3001.5	49.4	958	15	US-10-192-440-2	Sequence 9, Appl1
15	2982	49.1	950	15	US-10-192-440-9	Sequence 9, Appl1

16	1636	26.9	1017	10	US-09-965-830-6	Sequence 6, Appl1
17	1568	25.8	1107	15	US-10-185-867-16	Sequence 16, Appl1
18	1551	25.5	1102	15	US-10-185-867-36	Sequence 1, Appl1
19	1541.5	25.4	1083	15	US-09-160-224-1	Sequence 2, Appl1
20	1540.5	25.3	1083	10	US-09-965-830-2	Sequence 2, Appl1
21	1540.5	25.3	1083	15	US-10-185-867-2	Sequence 20, Appl1
22	1536	25.3	1082	9	US-09-119-855-2	Sequence 4, Appl1
23	1529.5	25.2	1080	15	US-09-813-320-4	Sequence 11, Appl1
24	1515.5	24.9	521	10	US-09-813-320-2	Sequence 11, Appl1
25	1511.5	24.9	530	10	US-09-119-855-11	Sequence 9, Appl1
26	1485.5	24.4	1284	9	US-10-160-224-9	Sequence 11, Appl1
27	1485.5	24.4	1284	15	US-10-185-867-11	Sequence 9, Appl1
28	1485.5	24.4	1284	15	US-10-174-613-7	Sequence 21, Appl1
29	1453.5	23.3	1174	15	US-10-188-296-21	Sequence 21, Appl1
30	1419	23.3	960	15	US-10-188-296-21	Sequence 21, Appl1
31	1419	23.3	960	15	US-10-188-341-21	Sequence 21, Appl1
32	1419	23.3	960	15	US-10-188-341-21	Sequence 21, Appl1
33	1419	23.3	960	15	US-10-188-297-21	Sequence 21, Appl1
34	1415	23.3	962	15	US-10-188-308-3	Sequence 3, Appl1
35	1415	23.3	962	15	US-10-188-296-3	Sequence 3, Appl1
36	1415	23.3	962	15	US-10-188-341-3	Sequence 3, Appl1
37	1415	23.3	962	15	US-10-188-297-3	Sequence 3, Appl1
38	1415	23.3	962	15	US-10-174-613-5	Sequence 5, Appl1
39	1409.5	23.2	962	15	US-10-174-613-4	Sequence 4, Appl1
40	1407	23.1	962	15	US-10-188-308-24	Sequence 24, Appl1
41	1407	23.1	962	15	US-10-188-296-24	Sequence 24, Appl1
42	1407	23.1	962	15	US-10-188-341-24	Sequence 24, Appl1
43	1407	23.1	962	15	US-10-188-297-24	Sequence 22, Appl1
44	1406.5	23.1	987	15	US-10-188-308-22	Sequence 22, Appl1
45	1406.5	23.1	987	15	US-10-188-296-22	Sequence 22, Appl1

ALIGNMENTS

RESULT 1
US-09-735-995-2
; Sequence 2, Application US/09735995
; Patent No. US20010034024A1
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Splawski, Igor
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; TITLE OF INVENTION: SYNDROME GENE
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/735, 995
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/226, 012
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-995-2

Query Match 100.0%; Score 6079; DB 9; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MPVRGHVAPONTFLDTIIRKFGOSRKFIARVENCNAVINCNDGFCGLGYSRAEVA 60
DB 1 MPVRGHVAPONTFLDTIIRKFGOSRKFIARVENCNAVINCNDGFCGLGYSRAEVA 60
OY 61 ORPCTDFLHGPTORRAAQIQAOLIGAEEKVEIAFYRKDSCFLCLVDVVPVKNEDG 120
DB 61 ORPCTDFLHGPTORRAAQIQAOLIGAEEKVEIAFYRKDSCFLCLVDVVPVKNEDG 120
OY 121 AVTMFLNFEVWEKMDVSSPADTNRGRPTSWLAPGRAKTFRLLPALATARESSV 180
DB 121 AVTMFLNFEVWEKMDVSSPADTNRGRPTSWLAPGRAKTFRLLPALATARESSV 180

181 RSGGAGAGAGAVVVDLTPAAPSSSLADEVTAMDNVAGLGAEEERALVGPSP 240
181 RSGGAGAGAGAVVVDLTPAAPSSSLADEVTAMDNVAGLGAEEERALVGPSP 240
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVRRASSADDEARAGVLP 300
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVRRASSADDEARAGVLP 300
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDRE 360
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDRE 360
361 PKIERHTNTEKTYOVLISGADVLPEKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
361 PKIERHTNTEKTYOVLISGADVLPEKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
421 TAVFTPYSAALFKETEBGPATECGYACQPLAVVDLIVDIMEFYDILLNFRITYV 480
421 TAVFTPYSAALFKETEBGPATECGYACQPLAVVDLIVDIMEFYDILLNFRITYV 480
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVARK 540
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVARK 540
541 RYSEYGAAYFLMLCTPALIHMLACTIWAIGNNEQPMDSRIGLHMLGQIGKPYNS 600
541 RYSEYGAAYFLMLCTPALIHMLACTIWAIGNNEQPMDSRIGLHMLGQIGKPYNS 600
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIF 660
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIF 660
661 AIIOQLYSGARHYTQMLRVRETRFHQIPNPLRQRLREYFQOHAMSTYNGIDMAV 720
661 AIIOQLYSGARHYTQMLRVRETRFHQIPNPLRQRLREYFQOHAMSTYNGIDMAV 720
721 PECLOADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGD 780
721 PECLOADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGD 780
781 FISGSEILRGDVVAALIGKNDIFGEPLNLARPGKSGNDVRLATYCDLKHIRH 840
781 FISGSEILRGDVVAALIGKNDIFGEPLNLARPGKSGNDVRLATYCDLKHIRH 840
841 VLDNYPERSDHFWSLETFENLRDNTMIPEGSPGTELEGFSROKRLSFRRTD 900
841 VLDNYPERSDHFWSLETFENLRDNTMIPEGSPGTELEGFSROKRLSFRRTD 900
901 QPGEVSALGPRAGAGPSRGRPGPGWGESPPSSGSPSESEDEGPGSSSPLRL 960
901 QPGEVSALGPRAGAGPSRGRPGPGWGESPPSSGSPSESEDEGPGSSSPLRL 960
961 PRPGEPPGEPPLMEDEKSSDTCNPLSGAFSGVSNIFSGDGRGROYELPRCP 1020
961 PRPGEPPGEPPLMEDEKSSDTCNPLSGAFSGVSNIFSGDGRGROYELPRCP 1020
1021 SLNINPLSSPGRARRBGDESRLDALQROLNRLERLSDAMATVQLQOROM 1080
1021 SLNINPLSSPGRARRBGDESRLDALQROLNRLERLSDAMATVQLQOROM 1080
1081 VTTGCGPPTSTPLPSPPLPTLTLDSISQVSEMACELPPGABELPQEBPT 1140
1081 VTTGCGPPTSTPLPSPPLPTLTLDSISQVSEMACELPPGABELPQEBPT 1140
1141 OLGAULTSOPLHRHGSDEGS 1159
1141 OLGAULTSOPLHRHGSDEGS 1159

GENERAL INFORMATION:
APPLICANT: Keating, Mark T.
TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
FILE REFERENCE: 2323-136
CURRENT APPLICATION NUMBER: US/09/735, 995
PRIOR APPLICATION NUMBER: 09/226, 012
PRIOR FILING DATE: 1999-01-06
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1159
TYPE: PRT
ORGANISM: Homo sapiens
US-09-735-995-4
Query Match 100.0%; Score 6079; DB 9; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MPVRGHAPONTFLDTIRKEGOSRKFITANARVENCAYICNDGFCELGYSRAEVM 60
1 MPVRGHAPONTFLDTIRKEGOSRKFITANARVENCAYICNDGFCELGYSRAEVM 60
61 QRPCTCFLHGPRTORRAAQAQALGAEERKVEIAFYRRKSGCFLCLVDVPEVKNE 120
61 QRPCTCFLHGPRTORRAAQAQALGAEERKVEIAFYRRKSGCFLCLVDVPEVKNE 120
121 AVTFILNFEVMEKDMGSPAHDTNHRGPTSMAPRATFRKLKLPALLATLARESSV 180
121 AVTFILNFEVMEKDMGSPAHDTNHRGPTSMAPRATFRKLKLPALLATLARESSV 180
121 AVTFILNFEVMEKDMGSPAHDTNHRGPTSMAPRATFRKLKLPALLATLARESSV 180
181 RSGGAGAGAGAVVVDLTPAAPSSSLADEVTAMDNVAGLGAEEERALVGPSP 240
181 RSGGAGAGAGAVVVDLTPAAPSSSLADEVTAMDNVAGLGAEEERALVGPSP 240
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVRRASSADDEARAGVLP 300
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVRRASSADDEARAGVLP 300
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDRE 360
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDRE 360
361 PKIERHTNTEKTYOVLISGADVLPEKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
361 PKIERHTNTEKTYOVLISGADVLPEKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
421 TAVFTPYSAALFKETEBGPATECGYACQPLAVVDLIVDIMEFYDILLNFRITYV 480
421 TAVFTPYSAALFKETEBGPATECGYACQPLAVVDLIVDIMEFYDILLNFRITYV 480
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVARK 540
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVARK 540
541 RYSEYGAAYFLMLCTPALIHMLACTIWAIGNNEQPMDSRIGLHMLGQIGKPYNS 600
541 RYSEYGAAYFLMLCTPALIHMLACTIWAIGNNEQPMDSRIGLHMLGQIGKPYNS 600
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIF 660
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIF 660
661 AIIOQLYSGARHYTQMLRVRETRFHQIPNPLRQRLREYFQOHAMSTYNGIDMAV 720
661 AIIOQLYSGARHYTQMLRVRETRFHQIPNPLRQRLREYFQOHAMSTYNGIDMAV 720
721 PECLOADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGD 780
721 PECLOADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGD 780

RESULT 2
US-09-735-995-4
; Sequence 4, Application US/09735995
; Patent No. US20010034024A1

QY 781 FISRGSEILRGDVVVAIIIGKNDIFGEPLNLVAPGKSGNDVRLATYCDLHKIHRDLE 840
DB 781 FISRGSEILRGDVVVAIIIGKNDIFGEPLNLVAPGKSGNDVRLATYCDLHKIHRDLE 840
QY 841 VLDMPYEFSDHFWSSLEITFNLDRDTNMIKPGSGTELEGFSRQRRKLSFRRTDKTE 900
DB 841 VLDMPYEFSDHFWSSLEITFNLDRDTNMIKPGSGTELEGFSRQRRKLSFRRTDKTE 900
QY 901 QPGEVSLGPRGAGAGSSGRGPGWGESPPSSGESSEDEGPRSSSPLRVFSS 960
DB 901 QPGEVSLGPRGAGAGSSGRGPGWGESPPSSGESSEDEGPRSSSPLRVFSS 960
QY 961 PRPGEPPGEPPLMEDCEKSDTCNPLSGAFSGVSNFISFWGDSRGROYELPCAPTP 1020
DB 961 PRPGEPPGEPPLMEDCEKSDTCNPLSGAFSGVSNFISFWGDSRGROYELPCAPTP 1020
QY 1021 SLNIPILSSPGRRRGDSRDLALQRLNRLTSLADNATVQLLQRMOTLVPPAYSA 1080
DB 1021 SLNIPILSSPGRRRGDSRDLALQRLNRLTSLADNATVQLLQRMOTLVPPAYSA 1080
QY 1081 VTPPGPPTSTPLPSPPLTLTLDLSQVSOFMACEELPRGAPELPQGPTRRLSLPG 1140
DB 1081 VTPPGPPTSTPLPSPPLTLTLDLSQVSOFMACEELPRGAPELPQGPTRRLSLPG 1140
QY 1141 QLGALTSQPLHRHSGDPGS 1159
DB 1141 QLGALTSQPLHRHSGDPGS 1159

RESULT 3
US-09-119-855-10
; Sequence 10, Application US/09119855
; Patent No. US20020099197A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A.J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: nmi-055
; CURRENT APPLICATION NUMBER: US/09/119, 855
; CURRENT FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-119-855-10

Query Match 100.0%; Score 6079; DB 9; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVRGHVAPONTFLDTIIRKFEQSQRKFIANARVENCAYIYCNDGFCGLGYSRAEVM 60
DB 1 MPVRGHVAPONTFLDTIIRKFEQSQRKFIANARVENCAYIYCNDGFCGLGYSRAEVM 60
QY 61 QPCTCDELHGPRTORRAAQIAQALGAEERKEVETAFYRKDSGCELCVADVVPKNEDEG 120
DB 61 QPCTCDELHGPRTORRAAQIAQALGAEERKEVETAFYRKDSGCELCVADVVPKNEDEG 120
QY 121 AVIMEILNFEVVEKMDVWSPAHDTNHRGPTSMVAPGRATFRLKLPALLALTARSSV 180
DB 121 AVIMEILNFEVVEKMDVWSPAHDTNHRGPTSMVAPGRATFRLKLPALLALTARSSV 180
QY 181 RSGGAGAGAPGAVVDVDTLTPAAPSESLALDEVTAMDNHVAAGLPAEERRALVGPSP 240
DB 181 RSGGAGAGAPGAVVDVDTLTPAAPSESLALDEVTAMDNHVAAGLPAEERRALVGPSP 240
QY 241 PRSAPGOLPSRAHSLMPDASGSSCLAFTRRSRSCASVRRASSADDIEMRAGVLP PPP 300
DB 241 PRSAPGOLPSRAHSLMPDASGSSCLAFTRRSRSCASVRRASSADDIEMRAGVLP PPP 300
QY 301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDFLASPTSDREIIA 360

DB 301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDFLASPTSDREIIA 360
QY 361 PKIKERHNTKEVTOVLSIGADVLPYKIQOARIRHMTLHYSPFAVADWMLILLYIY 420
DB 361 PKIKERHNTKEVTOVLSIGADVLPYKIQOARIRHMTLHYSPFAVADWMLILLYIY 420
QY 421 TAVFTPSAFLKETEPEGPPATECGYACOPVAVDVLVDIMEVDTLINFRTTYVANE 480
DB 421 TAVFTPSAFLKETEPEGPPATECGYACOPVAVDVLVDIMEVDTLINFRTTYVANE 480
QY 481 EVVSHPGRIAVHYFKGFLIDMVAALPFDDLIFGSGSEELIGLKTARLLRLVAVARKLD 540
DB 481 EVVSHPGRIAVHYFKGFLIDMVAALPFDDLIFGSGSEELIGLKTARLLRLVAVARKLD 540
QY 541 RYSEYGAVALFELMCFTFALAHMLACTIWAIGNNEQPHMDSRGWLNHLDQIGKPNSS 600
DB 541 RYSEYGAVALFELMCFTFALAHMLACTIWAIGNNEQPHMDSRGWLNHLDQIGKPNSS 600
QY 601 GLGSPSIXDKYVYALYTFESSLTSVCGVNSPMTNSEKIFISICVMLIGSLMTASIFGNVS 660
DB 601 GLGSPSIXDKYVYALYTFESSLTSVCGVNSPMTNSEKIFISICVMLIGSLMTASIFGNVS 660
QY 661 AIIORLYSGTARYHTOMLRVREFTRHQIDNPLRQRLSEXYFOHAMSYTNGIDMANAVLKG 720
DB 661 AIIORLYSGTARYHTOMLRVREFTRHQIDNPLRQRLSEXYFOHAMSYTNGIDMANAVLKG 720
QY 721 PECLQADICLHLNLSLQHRKPPRGATGCLRALAMKFTTHAPPGTIVHAGDLTALY 780
DB 721 PECLQADICLHLNLSLQHRKPPRGATGCLRALAMKFTTHAPPGTIVHAGDLTALY 780
QY 781 FISRGSEILRGDVVVAIIIGKNDIFGEPLNLVAPGKSGNDVRLATYCDLHKIHRDLE 840
DB 781 FISRGSEILRGDVVVAIIIGKNDIFGEPLNLVAPGKSGNDVRLATYCDLHKIHRDLE 840
QY 841 VLDMPYEFSDHFWSSLEITFNLDRDTNMIKPGSGTELEGFSRQRRKLSFRRTDKTE 900
DB 841 VLDMPYEFSDHFWSSLEITFNLDRDTNMIKPGSGTELEGFSRQRRKLSFRRTDKTE 900
QY 901 QPGEVSLGPRGAGAGSSGRGPGWGESPPSSGESSEDEGPRSSSPLRVFSS 960
DB 901 QPGEVSLGPRGAGAGSSGRGPGWGESPPSSGESSEDEGPRSSSPLRVFSS 960
QY 961 PRPGEPPGEPPLMEDCEKSDTCNPLSGAFSGVSNFISFWGDSRGROYELPCAPTP 1020
DB 961 PRPGEPPGEPPLMEDCEKSDTCNPLSGAFSGVSNFISFWGDSRGROYELPCAPTP 1020
QY 1021 SLNIPILSSPGRRRGDSRDLALQRLNRLTSLADNATVQLLQRMOTLVPPAYSA 1080
DB 1021 SLNIPILSSPGRRRGDSRDLALQRLNRLTSLADNATVQLLQRMOTLVPPAYSA 1080
QY 1081 VTPPGPPTSTPLPSPPLTLTLDLSQVSOFMACEELPRGAPELPQGPTRRLSLPG 1140
DB 1081 VTPPGPPTSTPLPSPPLTLTLDLSQVSOFMACEELPRGAPELPQGPTRRLSLPG 1140
QY 1141 QLGALTSQPLHRHSGDPGS 1159
DB 1141 QLGALTSQPLHRHSGDPGS 1159

RESULT 4
US-10-000-151b-3
; Sequence 3, Application US/10000151B
; Publication No. US20030013136A1
; GENERAL INFORMATION:
; APPLICANT: Balser, Jeffrey R.
; TITLE OF INVENTION: HUMAN K_{CR1} REGULATION OF HERG POTASSIUM CHANNEL BLOCK
; FILE REFERENCE: Vanderbilt Ref No. US20030013136A1 VU0120; Attorney Docket No. US2
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In version 3.1

; SEQ ID NO 3
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
us-10-000-151b-3

```
Query Match      100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 MPVRGHAAPONTFDTIIRKEEGSRKFIITANAVENCAYIYNDGFCELGYSRAEVM 60
DB      1 MPVRGHAAPONTFDTIIRKEEGSRKFIITANAVENCAYIYNDGFCELGYSRAEVM 60
OY      61 ORPCTCDELHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNEDG 120
DB      61 ORPCTCDELHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNEDG 120
OY      121 AVIMFILNEVVMEDKMGSPAHDTNRHGPPTSMILAPGRAKTFRLKLPALLALTARESSV 180
DB      121 AVIMFILNEVVMEDKMGSPAHDTNRHGPPTSMILAPGRAKTFRLKLPALLALTARESSV 180
OY      181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTADNHNVAAGLGPAAEERRALVGPSP 240
DB      181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTADNHNVAAGLGPAAEERRALVGPSP 240
OY      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
DB      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
OY      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKDPPFLASTSPREITIA 360
DB      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKDPPFLASTSPREITIA 360
OY      361 PKIKERTHNTEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVITY 420
DB      361 PKIKERTHNTEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVITY 420
OY      421 TAVFTPTSAFLKTEEGPPATECGYACQPLAVVDLIVDMFIVDILINFRTYVANE 480
DB      421 TAVFTPTSAFLKTEEGPPATECGYACQPLAVVDLIVDMFIVDILINFRTYVANE 480
OY      481 EVVSHPGRIAVHYFGWFLDMVAIPBDLLIFGSGSEBELLGLKTARLLRYVARKLD 540
DB      481 EVVSHPGRIAVHYFGWFLDMVAIPBDLLIFGSGSEBELLGLKTARLLRYVARKLD 540
OY      541 RYSEYGAAVFLMCTEFALIAHMLACIYTAIGNMEQPMDSRIGLHMLGDOIGKPYNS 600
DB      541 RYSEYGAAVFLMCTEFALIAHMLACIYTAIGNMEQPMDSRIGLHMLGDOIGKPYNS 600
OY      601 GLGGBSIXDKYVVALYFTFSSLTSGFNVSPNTNSEKIFSIQVWLIGSLMYASIFGNV 660
DB      601 GLGGBSIXDKYVVALYFTFSSLTSGFNVSPNTNSEKIFSIQVWLIGSLMYASIFGNV 660
OY      661 AIIORLYSGTARYTHOMLRVREFIRFHOIIPNPLRQRLLEYFOHASTYNGIDMAVLCGF 720
DB      661 AIIORLYSGTARYTHOMLRVREFIRFHOIIPNPLRQRLLEYFOHASTYNGIDMAVLCGF 720
OY      721 PECLADICLHNSLLOHCKPFGATGCLRALAMKKTTHAPGCDLVNAGDILLALY 780
DB      721 PECLADICLHNSLLOHCKPFGATGCLRALAMKKTTHAPGCDLVNAGDILLALY 780
OY      781 FISGSEILRGDVVAVALLGKNDIFGEPLNYARPGKSGNDVRLATYCDLKHIRHDDLE 840
DB      781 FISGSEILRGDVVAVALLGKNDIFGEPLNYARPGKSGNDVRLATYCDLKHIRHDDLE 840
OY      841 VLDVYPERSDHFWSSLETTEMLRTNMTPGSPSTLEBGGESRQKRLSLRRRTDKYTE 900
DB      841 VLDVYPERSDHFWSSLETTEMLRTNMTPGSPSTLEBGGESRQKRLSLRRRTDKYTE 900
OY      901 QPGEVSALGPGRAGAPSSRGRPGGPMGESPSGSSPSESSEDEGPGSSSPLRLVPPSS 960
DB      901 QPGEVSALGPGRAGAPSSRGRPGGPMGESPSGSSPSESSEDEGPGSSSPLRLVPPSS 960
```

```
OY      961 PRPGEPPGGEPLMEDCKSSDTCNPLSGAPSGVSNIFSWGDSNGROYELPRCPAPTP 1020
DB      961 PRPGEPPGGEPLMEDCKSSDTCNPLSGAPSGVSNIFSWGDSNGROYELPRCPAPTP 1020
OY      1021 SLNITPLSSPGRRRPPGVEDYESRLDALQROLNRLERLSDAMATVLOLQOROTLVPPAYSA 1080
DB      1021 SLNITPLSSPGRRRPPGVEDYESRLDALQROLNRLERLSDAMATVLOLQOROTLVPPAYSA 1080
OY      1081 VTPPGPSTSPILPVPSPPLTLTLDLSQVSOFAWCEELPPGAEPLPOEGPTRRLSLPG 1140
DB      1081 VTPPGPSTSPILPVPSPPLTLTLDLSQVSOFAWCEELPPGAEPLPOEGPTRRLSLPG 1140
OY      1141 QLGALTSQPLHRHGSPPGS 1159
DB      1141 QLGALTSQPLHRHGSPPGS 1159

RESULT 5
US-10-193-692-5
; Sequence 5, Application US/10193692
; Publication No. US20030074682A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis Deutschland GmbH
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELEOST POTASS
; FILE REFERENCE: AR02-005C
; CURRENT APPLICATION NUMBER: US/10/193,692
; CURRENT FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
us-10-193-692-5

Query Match      100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 MPVRGHAAPONTFDTIIRKEEGSRKFIITANAVENCAYIYNDGFCELGYSRAEVM 60
DB      1 MPVRGHAAPONTFDTIIRKEEGSRKFIITANAVENCAYIYNDGFCELGYSRAEVM 60
OY      61 ORPCTCDELHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNEDG 120
DB      61 ORPCTCDELHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNEDG 120
OY      121 AVIMFILNEVVMEDKMGSPAHDTNRHGPPTSMILAPGRAKTFRLKLPALLALTARESSV 180
DB      121 AVIMFILNEVVMEDKMGSPAHDTNRHGPPTSMILAPGRAKTFRLKLPALLALTARESSV 180
OY      181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTADNHNVAAGLGPAAEERRALVGPSP 240
DB      181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTADNHNVAAGLGPAAEERRALVGPSP 240
OY      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
DB      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
OY      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKDPPFLASTSPREITIA 360
DB      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKDPPFLASTSPREITIA 360
OY      361 PKIKERTHNTEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVITY 420
DB      361 PKIKERTHNTEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVITY 420
OY      421 TAVFTPTSAFLKTEEGPPATECGYACQPLAVVDLIVDMFIVDILINFRTYVANE 480
DB      421 TAVFTPTSAFLKTEEGPPATECGYACQPLAVVDLIVDMFIVDILINFRTYVANE 480
OY      481 EVVSHPGRIAVHYFGWFLDMVAIPBDLLIFGSGSEBELLGLKTARLLRYVARKLD 540
```

```

|||||
Db 481 EVVSHPRIRIAVHYFKGFWLLIDMVAIIPFDLLIFSGSGSEELIGLTKTARLLRLVAVARKLD 540
QY 541 RYSEYGAAYFLIMCTFALIAHMLACIWAYAIGNNEQHMDSRIGMHLNLDQICKPYNSS 600
Db 541 RYSEYGAAYFLIMCTFALIAHMLACIWAYAIGNNEQHMDSRIGMHLNLDQICKPYNSS 600
QY 601 GLGSPSFKDKXYVALYTFESSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
Db 601 GLGSPSFKDKXYVALYTFESSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
QY 661 AIIQRLYSGTARHTOMLRVREFIRHQIPNPLRQRLSEYFQAHMSTYNGIDMNAVYKGF 720
Db 661 AIIQRLYSGTARHTOMLRVREFIRHQIPNPLRQRLSEYFQAHMSTYNGIDMNAVYKGF 720
QY 721 PECLQADICHLNLSLQHKKPRFGATKGCIRALAMFKTTTHAPPGDTLVHAGDILALY 780
Db 721 PECLQADICHLNLSLQHKKPRFGATKGCIRALAMFKTTTHAPPGDTLVHAGDILALY 780
QY 781 FISGSIIEILRGDVVAAILKNDIFGEPLMLYARPGKSGDVRLATYCDLHKIHRDLE 840
Db 781 FISGSIIEILRGDVVAAILKNDIFGEPLMLYARPGKSGDVRLATYCDLHKIHRDLE 840
QY 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKDTE 900
Db 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKDTE 900
QY 901 QPGEVSLGPRAGAGSSSGRPGGPGWGESPPSSPESSEDEGPRSSSPLRLVFPSS 960
Db 901 QPGEVSLGPRAGAGSSSGRPGGPGWGESPPSSPESSEDEGPRSSSPLRLVFPSS 960
QY 961 PRPGEPPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
Db 961 PRPGEPPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
QY 1021 SLNIPILSSGRRRGVDESRLDALQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080
Db 1021 SLNIPILSSGRRRGVDESRLDALQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080
QY 1081 VTPPGPSTSPILPVSPILTLTLDLSQVSOQMACBELPRGAPELPOGCPTRRLSLPG 1140
Db 1081 VTPPGPSTSPILPVSPILTLTLDLSQVSOQMACBELPRGAPELPOGCPTRRLSLPG 1140
QY 1141 QLGALTSQPLRHGSDPGS 1159
Db 1141 QLGALTSQPLRHGSDPGS 1159

RESULT 6
US-10-185-867-10
: Sequence 10, Application US/10185867
: Publication No. US20030104429A1
: GENERAL INFORMATION:
: APPLICANT: Curtis, Rory A.J.
: TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
: FILE REFERENCE: NMI-055CP
: CURRENT APPLICATION NUMBER: US/10/185, 867
: PRIOR FILING DATE: 2002-06-27
: PRIOR APPLICATION NUMBER: US/09/358, 383
: PRIOR FILING DATE: 1999-07-21
: PRIOR APPLICATION NUMBER: USSN 09/119, 855
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 1159
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-185-867-10

Query Match 100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MPVRGHVAPQNFELDTIIRKEGOSRKFITIANARVENCAVIYCNDFCELCGYSRAEVM 60
Db 1 MPVRGHVAPQNFELDTIIRKEGOSRKFITIANARVENCAVIYCNDFCELCGYSRAEVM 60
QY 61 QRPCTCFLHGPRTORRAAAQIAQALLGAEEKVEIAFYRKDSCFLVDVVPVKNEDG 120
Db 61 QRPCTCFLHGPRTORRAAAQIAQALLGAEEKVEIAFYRKDSCFLVDVVPVKNEDG 120
QY 121 AVIMFIINFEYWEKDWGSPAHDTNHRGPTSLAAGRAKTFELKLPALLALTARBSV 180
Db 121 AVIMFIINFEYWEKDWGSPAHDTNHRGPTSLAAGRAKTFELKLPALLALTARBSV 180
QY 181 RSGGAGAGAPGAVVVDLTPAPSSSESLALDVTAMDHNHAGLGAEEERALVGGSP 240
Db 181 RSGGAGAGAPGAVVVDLTPAPSSSESLALDVTAMDHNHAGLGAEEERALVGGSP 240
QY 241 PRSAPGOLPSPRAHSLNPDASGSSCLARTSRRESCASVRRASSADDIEAMRAGVLP 300
Db 241 PRSAPGOLPSPRAHSLNPDASGSSCLARTSRRESCASVRRASSADDIEAMRAGVLP 300
QY 301 RHASTGAMHPLRSGLSNSTSDSLVYRTISKIPQITLNFYDLKGFPLASPTSDEITIA 360
Db 301 RHASTGAMHPLRSGLSNSTSDSLVYRTISKIPQITLNFYDLKGFPLASPTSDEITIA 360
QY 361 PKIKERTHNTKTYOVLISGADVLPYKLOAPRIHMTILHVS PFKAVMDMLILLYI 420
Db 361 PKIKERTHNTKTYOVLISGADVLPYKLOAPRIHMTILHVS PFKAVMDMLILLYI 420
QY 421 TAVETPYSAFLKETEGBPATCEGYACOPLA VVDLIVDIMEVLINFRRTYVANE 480
Db 421 TAVETPYSAFLKETEGBPATCEGYACOPLA VVDLIVDIMEVLINFRRTYVANE 480
QY 481 EVVSHPRIRIAVHYFKGFWLLIDMVAIIPFDLLIFSGSGSEELIGLTKTARLLRLVAVARKLD 540
Db 481 EVVSHPRIRIAVHYFKGFWLLIDMVAIIPFDLLIFSGSGSEELIGLTKTARLLRLVAVARKLD 540
QY 541 RYSEYGAAYFLIMCTFALIAHMLACIWAYAIGNNEQHMDSRIGMHLNLDQICKPYNSS 600
Db 541 RYSEYGAAYFLIMCTFALIAHMLACIWAYAIGNNEQHMDSRIGMHLNLDQICKPYNSS 600
QY 601 GLGSPSFKDKXYVALYTFESSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
Db 601 GLGSPSFKDKXYVALYTFESSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
QY 661 AIIQRLYSGTARHTOMLRVREFIRHQIPNPLRQRLSEYFQAHMSTYNGIDMNAVYKGF 720
Db 661 AIIQRLYSGTARHTOMLRVREFIRHQIPNPLRQRLSEYFQAHMSTYNGIDMNAVYKGF 720
QY 721 PECLQADICHLNLSLQHKKPRFGATKGCIRALAMFKTTTHAPPGDTLVHAGDILALY 780
Db 721 PECLQADICHLNLSLQHKKPRFGATKGCIRALAMFKTTTHAPPGDTLVHAGDILALY 780
QY 781 FISGSIIEILRGDVVAAILKNDIFGEPLMLYARPGKSGDVRLATYCDLHKIHRDLE 840
Db 781 FISGSIIEILRGDVVAAILKNDIFGEPLMLYARPGKSGDVRLATYCDLHKIHRDLE 840
QY 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKDTE 900
Db 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKDTE 900
QY 901 QPGEVSLGPRAGAGSSSGRPGGPGWGESPPSSPESSEDEGPRSSSPLRLVFPSS 960
Db 901 QPGEVSLGPRAGAGSSSGRPGGPGWGESPPSSPESSEDEGPRSSSPLRLVFPSS 960
QY 961 PRPGEPPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
Db 961 PRPGEPPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
QY 1021 SLNIPILSSGRRRGVDESRLDALQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080
Db 1021 SLNIPILSSGRRRGVDESRLDALQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080

```

Query Match 100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-10-174-613-49

RESULT 7

US-10-174-613-49

; Sequence 49, Application US/10174613

; Publication No. US20030114354A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOG Y

; FILE REFERENCE: D0123 NP

; CURRENT FILING DATE: 2002-06-19

; PRIOR APPLICATION NUMBER: US 60/299,378

; PRIOR FILING DATE: 2001-06-19

; PRIOR APPLICATION NUMBER: US 60/300,614

; NUMBER OF SEQ ID NOS: 93

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 49

; LENGTH: 1159

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-174-613-49

Query Match 100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-10-174-613-49

RESULT 7

US-10-174-613-49

; Sequence 49, Application US/10174613

; Publication No. US20030114354A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL POTASSIUM CHANNEL WITH HOMOLOG Y

; FILE REFERENCE: D0123 NP

; CURRENT FILING DATE: 2002-06-19

; PRIOR APPLICATION NUMBER: US 60/299,378

; PRIOR FILING DATE: 2001-06-19

; PRIOR APPLICATION NUMBER: US 60/300,614

; NUMBER OF SEQ ID NOS: 93

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 49

; LENGTH: 1159

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-174-613-49

Query Match 57.5%; Score 3493.5; DB 15; Length 1177;
Best Local Similarity 61.1%; Pred. No. 1.2e-245;
Matches 757; Conservative 107; Mismatches 227; Indels 147; Gaps 32;

US-10-193-692-4

RESULT 8

US-10-193-692-4

; Sequence 4, Application US/10193692

; Publication No. US20030074682A1

; GENERAL INFORMATION:

; APPLICANT: Exelixis Deutschland GmbH

; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELBOST POTASS

; FILE REFERENCE: AR02-005C

; CURRENT APPLICATION NUMBER: US/10/193,692

; CURRENT FILING DATE: 2002-07-11

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1177

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: :

US-10-193-692-4

Query Match 57.5%; Score 3493.5; DB 15; Length 1177;
Best Local Similarity 61.1%; Pred. No. 1.2e-245;
Matches 757; Conservative 107; Mismatches 227; Indels 147; Gaps 32;

US-10-193-692-4

RESULT 8

US-10-193-692-4

; Sequence 4, Application US/10193692

; Publication No. US20030074682A1

; GENERAL INFORMATION:

; APPLICANT: Exelixis Deutschland GmbH

; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELBOST POTASS

; FILE REFERENCE: AR02-005C

; CURRENT APPLICATION NUMBER: US/10/193,692

; CURRENT FILING DATE: 2002-07-11

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1177

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: :

US-10-193-692-4

```

1092 SPLPV-----SPLPTLTLSLQVSO-----FWACELRPGAE---LPOBGP-- 1132
1015 SOLNRELETRKALINVLLOLRQMAVPYPAVSAYSDPLAHVPRAHPHSYLTAAHNT 1074
1047 ROLNRELETRLSADMATVLOLRQMTLVPPAYSAVT-----TPGPGPTS-----T 1091
Db 1015 SOLNRELETRKALINVLLOLRQMAVPYPAVSAYSDPLAHVPRAHPHSYLTAAHNT 1074
QY 1092 SPLPV-----SPLPTLTLSLQVSO-----FWACELRPGAE---LPOBGP-- 1132

```

```

Db 1075 TPLSLQITDASSPGKSPDVLSEKSPDLSLSSGILHTVASTDWTMSMSPTELEVPASGPLL 1134
QY 1133 -----TRRLSLPGQL---GALTSQP-LHRHGSPP 1157
Db 1135 QPPELLCSSLRFPSPDLPSLEPGTLEGGSPETQIRHVSOP 1172

RESULT 9
US-10-192-440-10
; Sequence 10, Application US/10192440
; Publication No. US20030082718A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; TITLE OF INVENTION: 52908, A HUMAN POTASSIUM CHANNEL, AND
; TITLE OF INVENTION: USFS, THEREOF
; FILE REFERENCE: MRI2001-009PIRMA
; CURRENT APPLICATION NUMBER: US/10/192,440
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/341,953
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/304,243
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1195
; TYPE: PRF
; ORGANISM: Rattus norvegicus
US-10-192-440-10

Query Match 54.5%; Score 3312.5; DB 15; Length 1195;
Best Local Similarity 57.5%; Pred. No. 2e-232;
Matches 724; Conservative 114; Mismatches 251; Indels 171; Gaps 31;

QY 1 MPVRGHVAPQNTFLDTIIRKFEQSKRTIIANARVENCAYIYNDGFCGLGYRAEVM 60
Db 1 MPVRGHVAPQNTFLDTIIRKFEQSKRTIIANARVENCAYIYNDGFCGLGYRAEVM 60
QY 61 QRPCTDDEFLGPRRTORRAAQIOAOLGAEERVELAFYKDCSPCLVDVVPVKNEDG 120
Db 61 QRPCTDDEFLGPRRTORRAAQIOAOLGAEERVELAFYKDCSPCLVDVVPVKNEDG 120
QY 121 AVIMFLINEFVNEKDMVSPRPAHDNHRGPTSMALPGAKTFRKLKPLALLATARESSV 180
Db 121 AVIMFLINEFVNEKDMVSPRPAHDNHRGPTSMALPGAKTFRKLKPLALLATARESSV 180
QY 121 VAMFPIINEFYVDEDEDAAP-----ERVNIPLVKSVNKKLGFKFRGRLVLTFRKQSL 175
Db 121 VAMFPIINEFYVDEDEDAAP-----ERVNIPLVKSVNKKLGFKFRGRLVLTFRKQSL 175
QY 181 RSGGAGAGAPGAVVVDVLTTPAAPSSSESLADEVTAMDNVAGLGPABERRALVGP-- 238
Db 176 PDEDF-----PDVVVID-----SSKHSDSVAMKHFKSPTKESCSPSEADDTKALIQSQ 225
QY 239 SPRRSAPGOL-----PSRAHSLNRPDASGSCSLARTRSRESCASVPRASADDIEMRAG 294
Db 226 SPLVNISGRPLDHSSPKRQMDRLYRDMLOSSQSLTHSSRSLSLCSIRKASSVHDIEGF--N 283
QY 295 VLPR-----PPRHAS-----TGAMHPLRSGLNSTSDSDVRYRTISKIPIQTLNFEYDLK 344
Db 284 VHPKNIFRDNHASEDNCRNKKGPRNHHKSSLSLSDSTSDSNLKNKSTINKIPIQTLNFSQDV 343
QY 345 GDFPLAS-PTSDELLIAPKIKERTHNTEKVTQVLSIGADVLEPYKLOAPRIHRTILHY 403
Db 344 TEKKNTSPSSDFTIILAPKVEKERTHNTEKVTQVLSIGADVLEPYKLOAPRIHRTILHY 403
QY 404 SPFKAWDMILLLVITYTAFTYSAAFLLKETEERGPRAEGCYAOPLAVNDVLIYDIF 463
Db 404 SPFKAWDMILLLVITYTAFTYSAAFLLKETEERGPRAEGCYAOPLAVNDVLIYDIF 462
QY 464 IVDLINFTTYVANEVSSHPRGIAVHFKGMFLIDMAAIPFDLLIFGSGSESE--L 520
Db 463 IDIILINFTTYVANEVSSHPRGIAVHFKGMFLIDMAAIPFDLLIFGSGSEDETTL 522
QY 521 IGLKTAARLRLVAVARKLDRISEYGAVALFLIMCTFALIAHMLACIWAIGMNEOPHMD 580
Db 523 IGLKTAARLRLVAVARKLDRISEYGAVALFLIMCTFALIAHMLACIWAIGMNEOPYLT 582

```


Matches 670; Conservative 85; Mismatches 153; Indels 329; Gaps 32;

```

OY 1 MVRRGHVAPONTFLDTIIRKFEKGSRKFTIANAVENCAYICNDGCELCGYSRAEYM 60
DB 1 MVRRGHVAPONTFLDTIIRKFEKGSRKFTIANAVENCAYICNDGCELCGYSRAEYM 60
OY 61 ORPCICDFLHGRTQRRAAQIAQALLGAEEKVEIATPRKQSGCFCLVDVVPKKNEDG 120
DB 61 ORPCICDFLHGRTQRRAAQIAQALLGAEEKVEIATPRKQSGCFCLVDVVPKKNEDG 120
OY 121 AVIMFLNEEVMKDMVGPSPADTNHRGPTSWLAPGRAKTFRLLKLPALLA-LTARESS 179
DB 121 AVIMFLNEEVMKDMVGPSPADTNHRGPTSWLAPGRAKTFRLLKLPALLA-LTARESS 179
OY 180 VR-----SGAGAGAGAGAVVVDVLPAPSSSLALDEVYAMDNNHVAAGLGAPEERRA 233
DB 180 VR-----SGAGAGAGAGAVVVDVLPAPSSSLALDEVYAMDNNHVAAGLGAPEERRA 233
OY 146 ORLLSOSFLGSESGHGRPG----- 164
DB 146 ORLLSOSFLGSESGHGRPG----- 164
OY 234 LVGPSPRRASAPGOLPSPRAHSLNPDASGSSCLARTSRSCASVRRASADTEAMRA 293
DB 234 LVGPSPRRASAPGOLPSPRAHSLNPDASGSSCLARTSRSCASVRRASADTEAMRA 293
OY 165 --GPG-----PG----- 169
DB 165 --GPG-----PG----- 169
OY 294 GVLPPRRHASTGAMHPLRSGLLNSTSDLVRYRTISKIPQITLNFVDLKGDPPLASPT 353
DB 294 GVLPPRRHASTGAMHPLRSGLLNSTSDLVRYRTISKIPQITLNFVDLKGDPPLASPT 353
OY 170 -----TGR-----GKRTISOIQPOTLNFVEFNLEKRRSST 201
DB 170 -----TGR-----GKRTISOIQPOTLNFVEFNLEKRRSST 201
OY 354 SDRETIAP-KIKERTHNTEKTVOLSLGADVLPREKLOAPRIHRWTILHYSPEKAVMD 412
DB 354 SDRETIAP-KIKERTHNTEKTVOLSLGADVLPREKLOAPRIHRWTILHYSPEKAVMD 412
OY 202 TEIEIAPHKVERONTTEKTVOLSLGADVLPREKLOAPRIHRWTILHYSPEKAVMD 261
DB 202 TEIEIAPHKVERONTTEKTVOLSLGADVLPREKLOAPRIHRWTILHYSPEKAVMD 261
OY 413 LILLVITYAVFTPSAFLKETEGRPATCGACOPAVVDLIVDIMEFVDILINR 472
DB 413 LILLVITYAVFTPSAFLKETEGRPATCGACOPAVVDLIVDIMEFVDILINR 472
OY 262 LILLVITYAVFTPSAFLKETEGRPATCGACOPAVVDLIVDIMEFVDILINR 320
DB 262 LILLVITYAVFTPSAFLKETEGRPATCGACOPAVVDLIVDIMEFVDILINR 320
OY 473 TTYVANEEVSHPRGRIAVHYFKGFLIDMVAIFEDLLIFGSGSEE--LIGLTKTARL 529
DB 473 TTYVANEEVSHPRGRIAVHYFKGFLIDMVAIFEDLLIFGSGSEE--LIGLTKTARL 529
OY 321 TTYVNTNEVSHPRRIAVHYFKGFLIDMVAIFEDLLIFGSGDETTLLIGLTKTARL 380
DB 321 TTYVNTNEVSHPRRIAVHYFKGFLIDMVAIFEDLLIFGSGDETTLLIGLTKTARL 380
OY 530 LRLVAVARKLDRYSEYGAALFELMCTFALLAHMLACIYAGNERPYLEHRIKIGMDSL 589
DB 530 LRLVAVARKLDRYSEYGAALFELMCTFALLAHMLACIYAGNERPYLEHRIKIGMDSL 589
OY 381 LRLVAVARKLDRYSEYGAALFELMCTFALLAHMLACIYAGNERPYLEHRIKIGMDSL 440
DB 381 LRLVAVARKLDRYSEYGAALFELMCTFALLAHMLACIYAGNERPYLEHRIKIGMDSL 440
OY 590 GJOICKRPNNSSG-LGSPSKDKYVATVATFESSLSVCGNVSPTMNSKIKISICVMILG 648
DB 590 GJOICKRPNNSSG-LGSPSKDKYVATVATFESSLSVCGNVSPTMNSKIKISICVMILG 648
OY 441 GJOICKRPNNSSG-LGSPSKDKYVATVATFESSLSVCGNVSPTMNSKIKISICVMILG 500
DB 441 GJOICKRPNNSSG-LGSPSKDKYVATVATFESSLSVCGNVSPTMNSKIKISICVMILG 500
OY 649 SLMTASIFGNVSAIIQRLYSGTARKYHTQMLRVREFTRFQIIPNPLRORLEEFQOAMAST 708
DB 649 SLMTASIFGNVSAIIQRLYSGTARKYHTQMLRVREFTRFQIIPNPLRORLEEFQOAMAST 708
OY 501 SLMTASIFGNVSAIIQRLYSGTARKYHTQMLRVREFTRFQIIPNPLRORLEEFQOAMAST 560
DB 501 SLMTASIFGNVSAIIQRLYSGTARKYHTQMLRVREFTRFQIIPNPLRORLEEFQOAMAST 560
OY 709 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 768
DB 709 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 768
OY 561 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 620
DB 561 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 620
OY 769 LVHADDLTALYFIRSGSEIIRGQVYVAIILKNDIFGEPLVLYAPKGSNDVATLYC 828
DB 769 LVHADDLTALYFIRSGSEIIRGQVYVAIILKNDIFGEPLVLYAPKGSNDVATLYC 828
OY 621 LVHADDLTALYFIRSGSEIIRGQVYVAIILKNDIFGEPLVLYAPKGSNDVATLYC 680
DB 621 LVHADDLTALYFIRSGSEIIRGQVYVAIILKNDIFGEPLVLYAPKGSNDVATLYC 680
OY 829 DLHKIHRDLEVLMPYEFSDHFWSSLETFNLDT-----NMLPGSGSELGSGFSR 883
DB 829 DLHKIHRDLEVLMPYEFSDHFWSSLETFNLDT-----NMLPGSGSELGSGFSR 883
OY 681 DLHKIHRDLEVLMPYEFSDHFWSSLETFNLDT-----NMLPGSGSELGSGFSR 738
DB 681 DLHKIHRDLEVLMPYEFSDHFWSSLETFNLDT-----NMLPGSGSELGSGFSR 738
OY 884 QRRKLSFRRTDKDTEQPGEVSAIAPGAGAGAPSSRGHFGPMGESPSGSS--SPSS 941
DB 884 QRRKLSFRRTDKDTEQPGEVSAIAPGAGAGAPSSRGHFGPMGESPSGSS--SPSS 941
OY 739 -----LSDNGSGSPHE--LGP-----QFPGKGYSLIAGPGSQ 767
DB 739 -----LSDNGSGSPHE--LGP-----QFPGKGYSLIAGPGSQ 767
OY 942 EDEGGRSSSPLRLVFPSSPRPPEPPEGEPRLMECCCKSSDTCPNLSCAFSVSNIFSEW 1001
DB 942 EDEGGRSSSPLRLVFPSSPRPPEPPEGEPRLMECCCKSSDTCPNLSCAFSVSNIFSEW 1001
OY 768 NSMGAG-----PCAPGHDAAPPL-----SISDASGIW 795
DB 768 NSMGAG-----PCAPGHDAAPPL-----SISDASGIW 795
OY 1002 GDSRGRYOELPRCAPPTPSLINTPLSSPGRPRGDVE-----SRDLAOROLRLET 1054
DB 1002 GDSRGRYOELPRCAPPTPSLINTPLSSPGRPRGDVE-----SRDLAOROLRLET 1054
OY 796 PE-----LLOEAP-----PRHSP--QSPQEDPOCPWLKLSRLQLOAOMNRLES 838
DB 796 PE-----LLOEAP-----PRHSP--QSPQEDPOCPWLKLSRLQLOAOMNRLES 838

```

```

OY 1055 RUSADMAVLOLLOROM-----TLVPAYSAVTPPG-----GPTS 1090
DB 839 RVSSDLSTRILQJQPMPOGHASYILAEAPASNDLALVPLA-SETTSPGRPLPQGLPAPQ 897
OY 1091 T-----SPILFVSPPLPITLDSLOVQFMACELPFG-----APELP 1128
DB 896 TFSYDLDLDCSPKRRNSPRMP--HIAVATDXTLAPSS--OQOPGLMPLASPLHP 951
OY 1129 OE-----GPTRLSLPGOLGALTSQ-PLHRRGSDPG 1158
DB 952 LEVQGLIGCPGRS-SLPEHILGSPKQIDQRRGSGPG 987

RESULT 13
US-10-325-430-21
; Sequence 21, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Siles-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; TITLE OF INVENTION: 32838, 336 and 52908
; FILE REFERENCE: MP101-294P1RM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: SE 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 958
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-325-430-21

Query Match 49.4%; Score 3001.5; DB 12; Length 958;
Best Local Similarity 53.7%; Pred. No. 6.8e-210;
Matches 659; Conservative 83; Mismatches 139; Indels 347; Gaps 30;

```



```

Db      766 -----MPPHSPOQEDP-----DC-----W--- 782
QY      1004 SNGROYELRCPAPPLSLNLTSSPGHRRPGDVESRLDALROLNRLERLSADMATV 1063
Db      783 -----PLK-----LGRLEFOLOAOQNMRLRESVSSDLSRT 811
QY      1064 LQLOLROM-----TLVPPAYSAVTPPGP-----GPTST----- 1091
Db      812 LQLOLROMPOGHASYLEAPASNDLALVPIA-SETTSPGPRLPQGFILPACQTPSYGDLDD 870
QY      1092 -----SPILPSPPLTLTLDLSOVSQFMACELEPPG-----APETPQE-----G 1131
Db      871 CSPKRRNSSPRMP--HLAVADTKTLPSE-----OEQEPGLMPPLASPLHLEVOGLICG 924
QY      1132 PTRRLSLPQOLGALTSQ--PLHRHGSDPG 1158
Db      925 PCFS-SLPEHLSGVKQDLFORHGSDPG 951

```

```

RESULT 15
US-10-192-440-9
; Sequence 9, Application US/10192440
; Publication No. US20030082718A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: 52908, A HUMAN POTASSIUM CHANNEL, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MP12001-009PIRNM
; CURRENT APPLICATION NUMBER: US/10/192,440
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/341,953
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/304,243
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-192-440-9

```

```

Query Match      49.1%; Score 2982; DB 15; Length 950;
Best Local Similarity 52.9%; Pred. No. 1.8e-208;
Matches 647; Conservative 87; Mismatches 146; Indels 344; Gaps 25;

```

```

QY      1 MPVRGHVAPONTFTDITIRKFEQSRKFTIANARVENCAVYCCNDGCELCGYSRAEYM 60
Db      1 MVRKGHVAPONTYTDITIRKFEQSRKFTLANAOMENCALTYCNDGCELCGYSRAEYM 60
QY      61 QRPCTDFLHGPRITQORRAAQIQAALLGAEEKVEIAFYRKDSCFLCLVDVVPYKNEDG 120
Db      61 QRPCTDFLHGPRITQORRAAQIQAALLGAEEKVEIAFYRKDSCFLCLVDVVPYKNEDG 120
QY      121 AVIMEILNEVYMEKQVNGSPAHDTNHRGPTSWLAPGRAKTERLKLPAALLATARESSV 180
Db      121 AVIMEILNEF-----DLAQLLAKSSS-----RSILQRIILSSFL----- 154
QY      161 RSGGAGGAGAPGAVVVDVLTTPAARSSSLADEVTAMDNHVAIGPRAERALVGPSP 240
Db      161 RSGGAGGAGAPGAVVVDVLTTPAARSSSLADEVTAMDNHVAIGPRAERALVGPSP 240
QY      241 PRSAGQQLSPRAHSLNDPASGSSCLARTSRSCASVRRASADIEAMRAGVLPPPP 300
Db      241 PRSAGQQLSPRAHSLNDPASGSSCLARTSRSCASVRRASADIEAMRAGVLPPPP 300
QY      174 ----- 173
QY      301 RHASTGAMHPLRSGLNSTSDSLVRYRTSKIPQITLNFVDLKGDPFLASPTSDREIJA 360
Db      174 -----KYRTVSOIQPQFTLNFVEFNLEKHRSGSTTEIETIA 208
QY      361 P-KIKERTHNTEKTYOVLISGADVLPEYKLOAPRIHKTIIHNSPFKAVWMDLILLVI 419
Db      361 P-KIKERTHNTEKTYOVLISGADVLPEYKLOAPRIHKTIIHNSPFKAVWMDLILLVI 419

```

```

Db      209 PAKVVERQNTKTYOVLISGADVLPEYKLOAPRIHKTIIHNSPFKAVWMDLILLVI 268
QY      420 YTAVFPTYSAAFLKETEGRPATECGYACOPLAVIDLVIDIMFIVDILINERTTYVAN 479
Db      269 YTAVFPTYSAAFLKETEGRPATECGYACOPLAVIDLVIDIMFIVDILINERTTYVAN 327
QY      480 EEVSHPGRIAVHFKGKFLIDMAVAIFPDLLIFGSSGE--LIGLKTARLRLVYA 536
Db      328 DEVVSHPRRIAVHFKGKFLIDMAVAIFPDLLIFGSSGE--LIGLKTARLRLVYA 387
QY      537 RKLDRYSEGAIVLELCTFALINHWLACIMYALGKNEOPMDSRIGMLHNLQOIGKP 596
Db      388 RKLDRYSEGAIVLELCTFALINHWLACIMYALGKNEOPMDSRIGMLHNLQOIGKP 447
QY      597 YNSGC-LGQPSIKDYVYALYFTFSSLSVGVGNVSPMTNSEKIFSIQVMIGSLMAYSI 655
Db      448 YNSGDPASGSPVQDKYVYALYFTFSSLSVGVGNVSPMTNSEKIFSIQVMIGSLMAYSI 507
QY      656 FGNVSAIIORLYSGTARHTQMLRVREIRRHQIPLNPLRQRLSEYFOHAMSTYNGIDMA 715
Db      508 FGNVSAIIORLYSGTARHTQMLRVREIRRHQIPLNPLRQRLSEYFOHAMSTYNGIDMA 567
QY      716 VLKGFPECLQADICHLNRSLQHCPRGATKGCRLALAMKFKTTNAPPGDTIVHAGDL 775
Db      568 VLKGFPECLQADICHLNRSLQHCPRGATKGCRLALAMKFKTTNAPPGDTIVHAGDL 627
QY      776 LTAIFYIRSGSIELRGDVVAALGKNDIFGEPLMLYARPGKSGNDVATLYCDLHKIHR 835
Db      628 LSTLYFIRSGSIELRGDVVAALGKNDIFGEPLMLYARPGKSGNDVATLYCDLHKIHR 687
QY      836 DDLEVLDMYRPFESCHFWSSLEITFNLDQTN---MIP-GSPGSTLELGGFSRQKRKLS 890
Db      688 ADLEVLDMYRPAFAOTFNKLEVFENLDAQGLQSTFROAPRGHDDPGGFF----- 738
QY      891 FRRRTDXTDTEQGEVSALGPGRAGAPSSRGRPGGPMWSESSSPSSSEDEBPGRSS 950
Db      739 -----LNDSSGAPSS----- 749
QY      951 SPLRLVFPSSPRPGEPPGEPGLMEDCEKSDPTCNPLSGAFSGVSNIFSFMGDSNGROYQ 1010
Db      750 -----PELLOQMPSPSP--NPRQDLDCWHRELGFKLQLOAQNMRLRESVSSDLSR 759
QY      1011 ELPRCPAPPLSL-NIPLSSPGRRRGVDES-----RLDALQOLNRLERLSADMAT 1062
Db      760 -----PELLOQMPSPSP--NPRQDLDCWHRELGFKLQLOAQNMRLRESVSSDLSR 808
QY      1063 VLQLOLROMTLVPPAY-----SAVTPPGGP-----TSTPPL----- 1095
Db      809 ILQLOLROMTLVPPAY-----SAVTPPGGP-----TSTPPL----- 1095
QY      1096 -----PVSPPLTLTLDLSOVSQFMACELEPPGAPBELPQEGPTRLSLPQOLGA-- 1144
Db      867 HIQTRNRPSPRPHVAMAMDTLVPS-----EQEPGGLSLPLSLPLLEVPGLGGSRF 922
QY      1145 -----LHSQP-----LHRHGSDPG 1158
Db      923 SLPEHLSGVKQDLFORHGSDPG 946

```

Search completed: September 23, 2003, 17:48:51
 Job time : 463 secs